Solution Structure of the Two Core Elements of Group II Introns and Their Interplay with Metal Ions

DISSERTATION
zur
Erlangung der naturwissenschaftlichen Doktorwürde
(Dr. sc. nat)

vorgelegt der
Mathematisch-naturwissenschaftlichen Fakultät
der
Universität Zürich

von
Maria Pechlaner
aus
Österreich

Promotionskomitee
Prof. Dr. Roland K. O. Sigel (Vorsitz und Leitung der Dissertation)
Prof. Dr. Oliver Zerbe
Prof. Dr. Jörg S. Hartig

Zürich 2013

Contents

Acknowledgements

Special thanks goes to my Ph.D. advisor Prof. Dr. Roland K. O. Sigel for welcoming me in his group and giving me the chance to dive into a completely new and fascinating field of research. I am deeply grateful for the support and encouragement throughout the years.

Thank you also to the members of my Ph.D. committee Prof. Dr. Oliver Zerbe and Prof. Dr. Jörg S. Hartig for their time and helpful suggestions and for kindly acting as referees.

I want to thank also all the other people that contributed to this thesis, both in scientific and other ways:

Daniela Donghi for the fruitful collaboration in the D1kz project. However, even more I want to thank you for your invaluable enthusiasm, unending patience, unfailing memory, scrutiny and cheerful support in all matters.

Silke Johannsen for guiding my first steps in the field of NMR and RNA structure calculation. Thank you in particular also for your sincere interest in all my problems and many helpful comments.

Miriam Skilandat for the good company in our shared office, for the open ear for all kinds of questions and for numerous sharp and funny comments that made me laugh so many times.

Daniela Donghi, Silke Johannsen and Miriam Skilandat for carefully reading and correcting this thesis.

Maximiliane Korth, Joachim Schnabel, Jens Loebus and Pallavi Choudhary who accompanied me from the beginning, but also to all

other past and present members of the Sigel-Freisinger-group that have made my time at the University of Zurich a happy and memorable one.

Susann Paulus for the initial help and instructions in the wet lab.

Sofia Gallo, Tamara Huber and Joachim Schnabel for the happy times at winter and summer group activities that you made possible by taking the efforts of organizing them.

Veronika Zelenay, Cinzia Finazzo and Bernd Knobloch who were the ones who started the AvD5 and D1kz projects, respectively.

Manfred Jöhri for the friendly support in all computer hardware and printing matters. Simon Jurt, Nadja Bross and Prof. Dr. Oliver Zerbe for maintenance and trouble-shooting at the NMR spectrometers. And in particular also to Dr. Helena Kovacs from Bruker Biospin for her interest and active support in NMR questions.

My family, my parents Anna and Christoph, my brother Raimund and my sister Agnes. I know I can always depend on you.

List of Abbreviations

A, Ade	Adenine
AP	2-aminopurine
AvD5	Domain 5 from Av.I5
AvD5-AP24	AvD5 with a 2-aminopurine at position 24
Av.I5	*Azotobacter vinelandii* Intron 5
COSY	correlation spectroscopy
C, Cyt	Cytosine
D1 to D6	Domain 1 to 6 of group II introns
DNA	deoxy-ribonucleic acid
G, Gua	Guanine
HSQC	heteronuclear single quantum coherence
IRE	Iron Responsive Element
NMR	Nuclear magnetic resonance
NOESY	nuclear Overhauser spectroscopy
NTP	nucleotide triphosphate
OiD5	Domain 5 from the group IIC intron from *O. iheyensis*
PAGE	polyacrylamide gel electrophoresis
PlD5	Domain 5 from Pl.lsu/2
Pl.lsu/2	*Pylaiella littoralis* group II intron found in RNA of the large ribosomal subunit
ppm	parts per million
Pur	Purine
Pyr	Pyrimidine
RMSD	Root mean square deviation
RNA	ribonucleic acid

ScD5	Domain 5 from Sc.ai5γ
Sc.ai5γ	*Saccharomyces cerevisiae* group II intron located in the mitochondrial cytochrome oxidase 1 gene
snRNA	small nuclear RNA
T, Thy	Thymine
TOCSY	total correlation spectroscopy
Tris	2-amino-2-hydroxymethyl-propane-1,3-diol
U6 ISL	U6 intramolecular stemloop
U, Ura	Uracil
W-C	Watson-Crick

1

Introduction

In the central dogma of molecular biology[1,2] RNA has its place as the conveyor of genetic information between DNA and proteins. However, RNA has many more roles in the cell apart from that of a passive messenger, as indicated by the tiny fraction of transcribed RNA that is actually translated into proteins[3]. Today the so-called non-coding RNAs (ncRNA) have been implicated in all steps of gene expression and beyond. In complex with proteins or on their own they are involved in RNA processing (snRNAs of the spliceosome, snoRNAs, RNase P , group I and II introns), protein synthesis (rRNAs in the ribosome, tRNA), protein recognition and transport (signal recognition particle, SRP) and they regulate gene expression on the transcriptional and translational level (riboswitches) up to the miRNA-induced transcript degradation[4]. As retrotransposons, mobile RNA elements like Alu and L1 are reverse transcribed and inserted back into DNA and thus significantly contribute to the modulation of modern genomes[5].

The discovery of catalytically active RNA[6,7], i.e. ribonucleic acid enzymes or ribozymes, raised many questions. Which factors govern the folding of those large polyanions into compact structures? With a — compared to the amino acids of proteins — very limited set of bulky building blocks, none of which have pK_as in the neutral range, where does the catalytic propensity come from? An important role of metal ions in both structure and catalytic activity was soon evident and the two things have often been found inseparably connected[8,9]. In addition, metal ions but also the precisely aligned functional groups in active sites of ribozymes can activate nucleobases and enable them to participate directly in catalysis (e.g. by acid-base catalysis through strongly shifted nucleobase pK_as).

With a growing body of RNA structures available in databases, we are at the exciting point where general principles of RNA architecture become observable[10]. However, in addition to the static pictures from crystal structures, especially the characterization of dynamic equilibria in solution, governed largely by the interplay of RNA with its physiological metal ion co-factors, will lead to a deeper understanding of their function.

1.1 Nucleic acid structure

1.1.1 Basic building blocks

Nucleic acids are built from nucleotides, composed of a sugar moiety, a 5'-phosphate group and one of four different nucleobases: the purines adenine and guanine and the pyrimidines cytosine and uracil (RNA) or thymine (DNA) (Figure 1.1). In RNA the sugar is a ribose, in DNA a 2'-deoxy-ribose. *In vivo*, DNA and RNA polymerases use nucleoside triphosphates as substrates to produce RNA and DNA polymers with 3'→5' phosphodiester linkages and one phosphate — and therefore one negative charge — per monomer. The robust viral T7 RNA polymerase performs well also *in vitro*, a property which is widely taken advantage of to produce RNA sequences, even if there are some sequence and length limitations[11,12]. The phosphoramidite-based chemical synthesis is much more costly, but allows the site-specific incorporation of isotopically labelled or modified nucleotides.

At physiological pH the nucleobases in free nucleotides are neutral[13]. The pK_as can, however, be significantly perturbed in higher order structures[14,15], where in particular A-N1 and C-N3 can also be found in their protonated form.

The conformation of the nucleotides is characterised by the sugar pucker and the glycosidic angle χ (Figure 1.1c,d). In addition, the six torsion angles of the sugar-phosphate backbone give rise to a larger number of rotamers (Figure 1.1a). Hydrogen bonding interactions of the nucleobases are often described by referring to one of three edges: the Watson-Crick (W-C) edge, the Hoogsteen edge and the sugar edge (Figure 1.1b). For instance, the interaction of A and U and

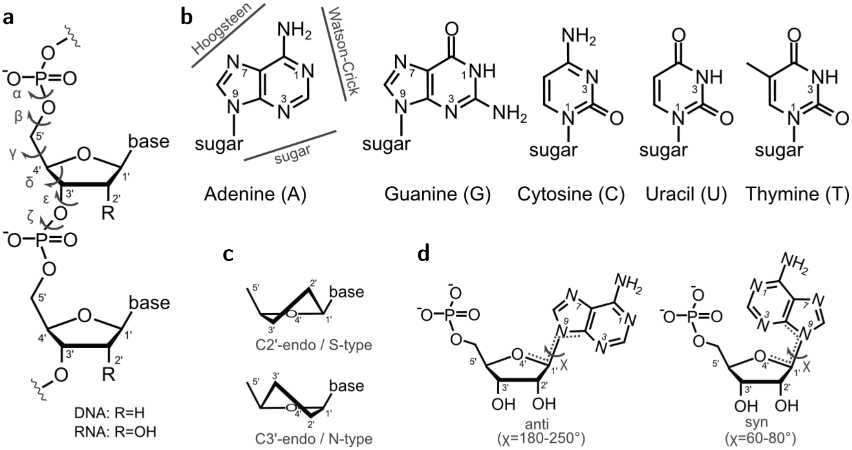


Figure 1.1 The basic building blocks of nucleic acids. a) Section of the sugar-phosphate backbone. The nucleotides are linked by 3'→5' phosphodiester linkages. The six torsional angles of the backbone are indicated. b) The five nucleobases occuring in RNA and DNA. c) The main sugar conformations in regular A- and B-form helices are C2'-endo and C3'-endo, respectively. d) Nucleobase orientations are classified as *anti* or *syn* according to the glycosidic angle χ.

of G and C via their Watson-Crick faces results in the canonical base pairs that are the basic unit of nucleic acid secondary structures (Figure 1.2a,b). They allow two complementary strands to anneal in an anti-parallel fashion yielding the well-known right-handed helical structures. The 2'OH group distinguishes RNA from DNA and renders it more susceptible to self-cleavage through in-line nucleophilic attack on the adjacent 3'-phosphate group. It is also responsible for the preference of A-form helical conformation with deep and narrow major groove and shallow minor groove and C3'-endo sugar conformation in RNA, whereas DNA usually adopts the B-form, with two grooves of similar depth and typically C2'-endo sugar conformation. Helical parameters in addition depend on sequence and external conditions[16,17].

1.1.2 *Principles of RNA tertiary structure*

DNA occurs mainly in its double-stranded, helical form. The variety of RNA structures, on the other hand, is determined by terminal and internal loops, bulges and multibranched junctions (Figure 1.3a-e) which interrupt standard helical regions. The structures of these non-helical regions are stabilized by

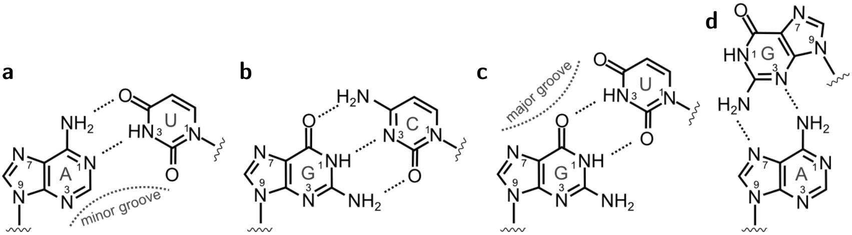


Figure 1.2 Common base pairs in RNA. a) and b)Watson-Crick base pairs. c) G-U wobble pair. d) Sheared G-A pair. The location of minor and major grooves is indicated in two examples.

hydrogen bonding and stacking interactions. They can introduce sharp turns, provide binding sites for proteins or other molecules and mediate tertiary interactions which determine the global structure of the RNA molecule.

Tertiary interactions can involve canonical base-pairing like in pseudoknots and kissing loops (Figure 1.3f,g) or other sequence-dependent contacts like tetraloop receptor interactions or base triples and quadruples. The 2'OH groups play an important role as hydrogen bond donors and acceptors in many tertiary interactions, among which is also the sequence-independent ribose zipper motif that holds together two RNA backbones[18].

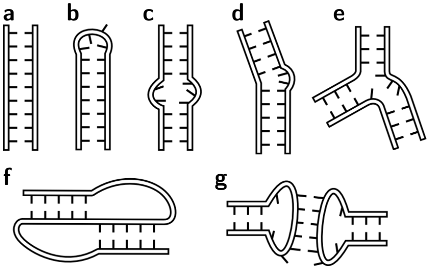


Figure 1.3 Apart from canonical double-stranded helices (a), RNA secondary structure contains terminal and internal loops (b,c), bulges (d), multibranch junctions (e), pseudoknots (f) and kissing loops (g).

Non-canonical base pairs and mismatches. While the G-U wobble pair is a very common component of RNA helices (Figure 1.2c), many other non-canonical base pairs or mismatches[19–22] (e.g. the sheared G-A pair (Figure 1.2d), G-A imino, A-U reverse Hoogsteen) are found more often outside the helical regions, in loops or bulges or in company with other mismatches, because they introduce steric perturbations in helices. Some of them require protonation at adenine N1 or cytosine N3 sites. Non-canonical base pairs are also part of base triples[23] and quadruples, thereby establishing tertiary contacts.

A-minor interactions. The smooth shape of the adenine base with its unobtrusive H2 group makes an interaction with the minor groove of helical regions particularly favourable. Such A-minor contacts, which are characterized by the interactions of the 2’OH groups, have been found to be one of the most common tertiary interactions[24,25]. They belong to one of four classes (Figure 1.4a) which frequently occur in clusters where the A-minor interaction with the higher type number is found 5’ of the one with the lower type number (i.e. if the first one is type III, the following one will be type II)[25].

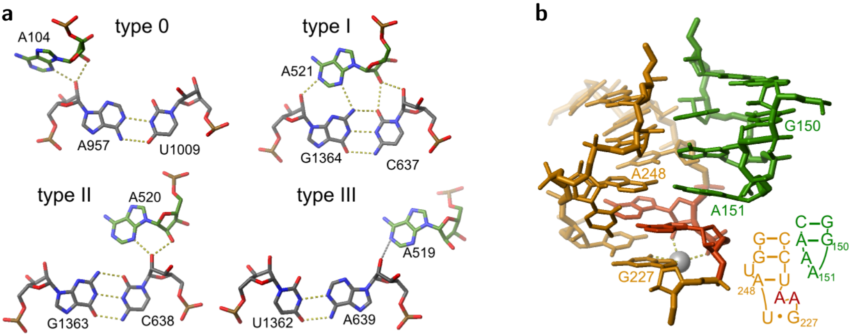


Figure 1.4 Two common tertiary interaction motifs. a) The four different types of A-minor interactions defined by Nissen et al.[25]. The picture was reproduced according to Nissen et al.[25] using the crystal structure of the large ribosomal subunit from *H.marismortui* (PDB-ID: 1FFK[26]). b) A GAAA tetraloop (green) bound to an 11-nt receptor (yellow) in the crystal structure of the P4-P6 domain of the Tetrahymena group I intron (PDB-ID: 1GID[27]). The motif contains an A-A platform (red) which is involved in coordination of a K(I) ion[28].

Tetraloops and receptors. Hairpins are most frequently capped by tetraloops, often belonging to one of two conserved folds[29,30]. While UNCG loops are very stable and inert closing structures[29], the GNRA loops interact with a variety of receptors[31,32] (where N is any nucleotide and R is a purine). The most complex and conserved type of receptor specific for GAAA tetraloops is an 11-nt motif involving an A-A platform and a network of hydrogen bonding interactions, but also simple tandem G-C base pairs can serve as receptors[33]. In both cases the adenines from the loop form A-minor interactions. The GNRA tetraloop is a surprisingly stable motif that can tolerate strand nicks and insertions inside its sequence and even deletion of the second nucleotide without losing the ability to bind to its receptor[34].

Stacking of the bases. The maximization of stacking interactions is believed to be one of the driving forces of RNA folding[10,35,36]. Even 3' overhanging nucleotides which tend to stack on the closing base pair can contributes significantly to stability[37,38]. Especially the co-axial stacking of helices in multibranch junctions will strongly affect the global fold and it has been shown to be influenced by di- or multivalent ions[39].

1.1.3 Dynamics

Nucleic acids in solution experience dynamics on a wide timescale, from the rapid interconversion of sugar puckers, stacking and unstacking of unpaired bases, base flipping and helix bending and twisting motions to changes in secondary structure (see recent reviews[40,41]). In many cases dynamic transitions are intrinsic to the function of the molecule like the catalysis of ribozymes or the regulatory function of a riboswitch[40,42]. The unpaired residues in loops and bulges often experience more varied and higher amplitude dynamics in the pico- to nanosecond range than residues in helices[41,43–45]. An increased mobility in these regions might be the basis for an adaptive recognition of binding partners[46,47].

1.1.4 RNA structure determination by nuclear magnetic resonance spectroscopy

In our attempt to elucidate the structural and dynamic basis of RNA function, solution NMR spectroscopy provides us with many valuable tools[41,48,49]. While the size limit — although in exceptional cases up to 100 kDa particles have been studied[50] — usually restricts NMR methods to rather small structures, they are especially powerful when it comes to the dynamic changes induced by small ligands, metal ions (see Section 1.2.1) and other co-factors on a per-atom-basis. NMR allows the characterization of transient interactions and minor conformational states, as well as subtle effects like the protonation at an unusual site, which is often difficult to detect in crystal structures.

Several comprehensive reviews have focused on nucleic acid structure determination by NMR[48,49,51–54], drawing a picture of the challenges, which are in many ways distinct from those of protein NMR. The nucleotide building blocks are comparatively large, which extends also to the size of basic structural elements and thus RNA molecules too easily exhaust the size limit of solution NMR. The nucleobases are characterized by a low proton density, which is a disadvantage since the ^{1}H resonances are the main source of information. The four chemically very similar nucleotides lead to relatively bad signal dispersion and concomitant strong overlap not only in the ^{1}H resonances, but also in the resonances of the heteronuclei ^{31}P, ^{13}C and ^{15}N. The naturally most abundant isotopes of carbon and nitrogen, ^{12}C and ^{14}N have spins 0 and 1, respectively, and we therefore use ^{13}C,^{15}N-enriched nucleotides for heteronuclear experiments. The signal overlap in the ^{1}H dimension can be reduced by the deuteration of selected proton positions, which in addition also increases relaxation times because protons are a major source of dipolar relaxation[55]. Another strategy to reduce overlap is the use of nucleotide-specific or segmental labelling in combination with heteronuclear-filtered experiments, which are designed to discriminate against ^{1}H spins attached to ^{13}C or ^{12}C (or ^{15}N and ^{14}N)[56]. Lastly, also three-dimensional experiments help to alleviate the overlap.

Sequential assignment. While several standard experiments can connect the amino acids in proteins via scalar couplings through the covalent bonds of the backbone, establishment of through-backbone connections across the phosphates in nucleic acids is more difficult. Therefore assignment usually depends more strongly on through-space (i.e. Nuclear Overhauser Enhancements, NOEs), rather than through-bond connections[57]. In helical regions there is a characteristic pattern of cross peaks connecting H1' to the aromatic H6/H8 of its own and the following nucleobase, allowing to assign a polynucleotide strand step by step, which is usually referred to as the "sequential walk" (Figure 1.5). Similar sequential connections are also possible for H2' and H3' to H6/H8. Adenine H2 are the only non-exchangeable protons on the Watson-Crick side, which is the side pointed away from the own sugar-phosphate backbone and can therefore provide especially valuable inter-strand connections. The imino protons exchange rapidly with the solvent and can only be observed when engaged in stable hydrogen bonds, usually in the W-C base pairs of the helices. Scalar

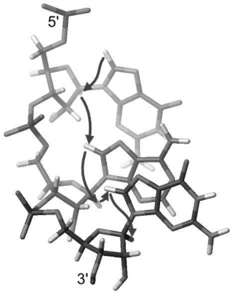


Figure 1.5 Sequential walk connectivities that can be observed via NOEs.

couplings from the proton to the donor and the acceptor in a hydrogen bond are commonly used to confirm secondary structure[58,59].

Dihedral angles. A variety of experiments are available to determine backbone and sugar torsion angles from scalar couplings (described e.g. in[49]). A qualitative evaluation of sugar puckers is also possible in ^{1}H,^{1}H-TOCSY experiments: C2'-endo conformation, which is most common in B-form helices, is characterized by $^3J_{H1',H2'}$ couplings of >8 Hz, leading to intense H1'-H2' and observable H1'-H3' peaks, which are absent in C3'-endo conformation, the dominant conformation in A-form helices. The N-glycosidic angle (χ) almost exclusively determines the intranucleotide H1'-H6/H8 distance (Figure 1.1d), therefore a very short (2.5 Å) distance is a strong indication for *syn*-conformation.

Residual dipolar couplings. NOEs yield only short-range distance restraints (<6 Å). In the usually rather elongated nucleic acid molecules where long-range NOEs between distant structural elements are rare or absent, they can not efficiently define the global bending. This information can be gained by measuring residual dipolar couplings (RDCs)[60]. In solution, the isotropic distribution of orientations usually averages the otherwise large dipolar couplings to zero. Already a very slight partial alignment of the molecules makes RDCs of a few Hz observable. Alignment of nucleic acids is usually achieved by addition of negatively charged filamentous phages[61,62]. The RDCs carry information on the orientation of the corresponding bond in the frame of the molecule. This provides long-range information that can be incorporated together with the NOE-derived distance information and determined dihedral angles into restrained molecular dynamics simulations to determine a structural model.

1.2 Metal ions in RNA structure and function

Nucleic acids are polyanions with one negative charge per nucleotide and naturally attract cations. Metal ions allow them to fold into compact structures despite the charged sugar-phosphate backbone[63], but also play an important role in the fine-tuning and diversification of chemical properties of the otherwise very homogeneous nucleotide building blocks[64–66].

Inside the cell, the most abundant metal ions are the group I and group II ions

K(I) and Mg(II) and they are expected to be the major interaction partners of RNA, even if also smaller amounts of other metal ions can have a surprisingly strong impact[67,68]. Intracellular activities of K(I) and Mg(II) are comparable to that of a solution containing 0.5-1.0 mM Mg(II) and 0.15 M monovalent K(I)[69–72]. Since both K(I) and Mg(II) are spectroscopically silent, metal ion–RNA interactions have often been studied using substitutes like Mn(II)[73], Cd(II)[74–76], cobalt(III)hexammine[77–81], Tb(III)[82,83] for Mg(II) or Tl(I) and NH_4^+ for K(I)[28,84]. However, it has to be kept in mind that the stronger polarizability of transition metal ions like Mn(II) and Cd(II) makes them much more prone to interact directly with the nucleobases, especially purine N7[85,86] than the closed-shell ions Mg(II) and K(I).

RNA stability is as much determined by a diffuse ion atmosphere as by site-specifically bound metal ions[86–89]. The diffuse ions interact with the electrostatic field of RNA and play an important role as non-specific charge screeners as only a fraction of negative phosphate charges is neutralized by site-bound metal ions that are observable in x-ray crystal structures[90,91]. The site-bound metal ions do not necessarily make specific inner-sphere contacts to RNA moieties. In fact, RNAs contain sites with a high density of electronegative groups that attract metal ions with low selectivity, indulging different metal ions and different binding modes[88,92–95]. The deep major groove of RNA is one such site, especially at the N7 and O6 of guanines and in particular at G-U wobble pairs where the major groove is lined by electronegative groups[96–99] (Figure 1.2c).

Mg(II) has a special place in RNA structure and mechanism, not only because of its intracellular abundance. Its small size and high charge density make it optimally suited to tightly coordinate six oxygen ligands. While on the one hand this goes along with a particularly large hydration free energy, which has to be overcome to form inner-sphere contacts, it also makes Mg(II) unbeatably efficient in stabilizing sites of high electronegative density at the sugar-phosphate backbone of nucleic acids. Highly coordinated Mg(II) is observed in buried solvent-inaccessible binding pockets[100,102–104] an example of which is shown in Figure 1.6a. The non-bridging phosphate oxygens are by far the most favoured ligands[90,91]. O6 and N7 of purine and O4 of pyrimidine nucleobases are less common as inner-sphere ligands, usually only in addition to inner- or outer-sphere contacts to phosphate oxygens[86,90,91]. For example, a sharp turn of the

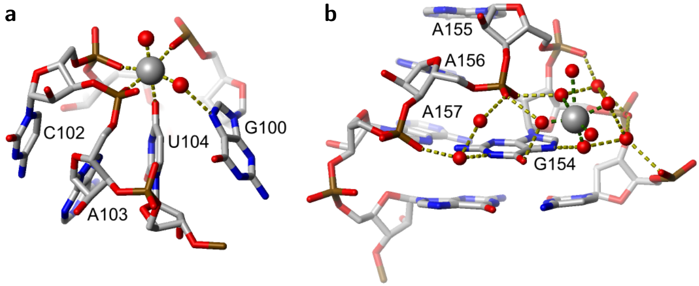


Figure 1.6 Examples of inner- and outer-sphere binding sites of Mg(II) in two crystal structures. a) Mg(II) makes four inner- and one outer-sphere (via a water molecule to G100N7) contact to phosphate and nucleobase ligands in the crystal structure of an M-box riboswitch[100] (PDB-ID: 2QBZ). b) Close to a GAAA tetraloop a hexahydrated Mg(II) is coordinated via a network of hydrogen bonds (adapted from Batey and Doudna[94], PDB-ID: 1DUL[101]). Water molecules are depicted as red spheres, hydrogen bonds as dashed yellow lines.

phosphate backbone like in the group II intron active site[105–107] (see also Section 1.4.1) can lead to the proximity of phosphate groups and can favour metal ion coordination. At other sites Mg(II) can keep its complete inner shell of water ligands but might still interact stably via a network of hydrogen bonds and ordered water molecules[94,108] (Figure 1.6b). Reports of affinities for directly coordinated metal ions in larger RNA structures are few because it is all but trivial to determine the affinity of a specifically chelated metal ion in the background of diffuse metal ions that are also required for a stable structure[109–111]. Leipply and Draper[112] recently determined an apparent intrinsic K_D of 2 µM for a known strong Mg(II) binding site in a 58mer rRNA. In smaller constructs, where deeply buried binding sites can not form, usually K_Ds in the millimolar range are observed[73,80,113–116].

K(I) is larger than Mg(II) with a lower charge density and much less efficient in neutralizing phosphate charges. Since two ions would be required to provide the same positive charge as one Mg(II) there is also an entropic disadvantage[117]. Nonetheless, there are also binding sites specific for monovalent metal ions[28,94,118,119] like the A-A platform that occurs in tetraloop receptors (Figure 1.4b).

Metal ions can be involved in ribozyme catalysis by stabilizing transition states and leaving groups, facilitating deprotonation and activating the nucleophile[9,65,66,120–122]. A two-metal-ion mechanism, where inner-sphere coordinated Mg(II) ions both activate the attacking nucleophile and stabilize the leaving group, has been proposed for the phosphoryl transfer reactions catalysed by the large group I and II ribozymes[107,123–125] (Section 1.3). However, direct coordination is not a prerequisite, metal ions can also influence the catalytic mechanism by longer range electrostatic interactions or water-mediated hydrogen bonds[64]. There is evidence that electronegative binding pockets, which are not strongly selective for a specific metal ion type but can bind a variety of different ions are nonetheless not restricted to structural roles but can directly be implicated in catalysis[64,95]. This could explain why the catalytic activity of small ribozymes does not depend strictly on a Mg(II), but can be promoted by a variety of different mono- and divalent metal ions[91,126–130].

1.2.1 Solution NMR spectroscopy to study RNA-metal ion interactions

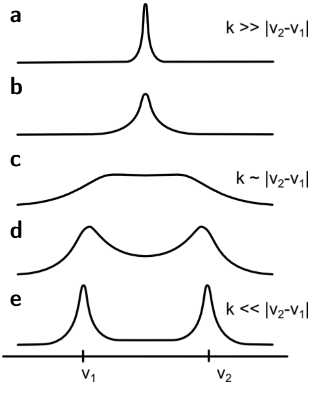


Figure 1.7 The influence of chemical exchange between two species on their NMR signals. a) Fast exchange leads to one averaged peak. As the exchange gets slower, peaks begin to broaden until disappearance (b,c), to reappear and sharpen at the separate resonances of the two species (d,e). Whether a process is fast or slow is determined by the exchange rates (k) relative to the frequency difference of the two species (ν_2-ν_1).

The binding of metal ions is reflected in several NMR observables of the RNA molecule: in chemical shifts and line widths, but in special cases also in scalar couplings and NOE connectivities[81,131–134]. Alternatively, binding can be monitored from the point of view of the metal ion resonance when an NMR active isotope is employed (^{23}Na(I), ^{39}K(I), and ^{25}Mg(II)), by which no information on binding sites but on the kinetics and thermodynamics of ion binding can be gained[135–139]. A quantitative analysis of line widths is possible in the intermediate exchange regime (Figure 1.7b,c) and can yield on- and off-rates of metal binding as well as equilibrium dissociation constants[140,141]. Cowan and colleagues found exchange rates in the range of 10^3 s^{-1} for weak binding sites in 5S rRNA and tRNA molecules[135,136], wich are comparable to the dissociation rate of Mg(II)-ATP/ADP complexes[142] and to the exchange rates in dsDNA, where outer-sphere Mg(II) coordination is suggested[137].

So far the observation of small direct couplings between metal ions and RNA moieties has been possible only in the exceptionally strong binding sites in a G-quadruplex[143], but also the indirect effect on the splitting of the ^{15}N signal as observed by $^{107/109}Ag$ in a modified DNA is a direct proof of coordination[133]. However, such strong interactions are an exception in small RNA molecules[81].

A majority of interactions is weaker and — in particular with the natural co-factor Mg(II) — involves an exchange rate that is intermediate to fast compared to the chemical shift difference in Hz[144], leading to one averaged signal at a position determined by the ratio of metal-bound:unbound molecules (Figure 1.7a,b). This ratio is gradually changed by titrating in the metal ion and the observed chemical shift perturbations mapped against metal ion concentration can yield apparent affinity constants[114,140] (see Equation 4.3 in the Experimental Section). When it is possible to assign the number and approximate affinity of binding sites in a molecule, these can serve as the starting point for an iterative procedure to determine the intrinsic affinity of each site[114,145]. However, the chemical shift perturbations can be caused by the approaching electron density of the metal ion as well as by induced structural rearrangements, two effects that can not easily be distinguished or untangled. For instance, a very slight shifting that moves a proton, which might be at some distance from the binding site, more into or out of the plain of an aromatic ring influences its chemical shift markedly. Only the very strong shift changes that are induced at direct coordination sites are unambiguous in that regard. However those can be measured straightforwardly only at the ^{15}N sites of the nucleobases, where upfield shifts of ~20 ppm at the N7 position are observed for Cd(II) in combination with a noticeable downfield shift of the adjacent C8 (2.3 ppm) and H8 (0.38 ppm)[74]. Even if the maximal upfield shift observed with Mg(II) in the same site is only 6.5 ppm[146], theoretical calculations indicate that

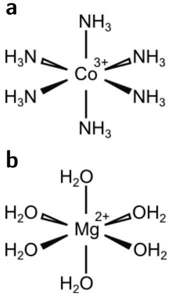


Figure 1.8 Cobalt(III)-hexammine (a) is commonly used as a mimic for $[Mg(H_2O)_6]^{2+}$ (b).

Mg(II) chelation can cause ^{15}N shifts similar to those of Cd(II)[147,148].

As the Mg(II) concentration is increased, the exchange of Mg(II) at low-specificity interaction sites leads to a general broadening of signals[142] (Figure 1.7b,c). Much stronger broadening is induced in a r^{-6} distance dependent way by the paramagnetic effect of a metal ion with an unpaired electron like Mn(II), which is often used as a probe to detect Mg(II) binding sites[149–151].

Another commonly used Mg(II) mimic is cobalt(III)hexammine ($[Co(NH_3)_6]^{3+}$), which has a similar geometry as $[Mg(H_2O)_6]^{2+}$[78] (Figure 1.8). Since its ligands are kinetically inert it is possible to observe NOEs from the ammonia protons to protons of the RNA, which is direct evidence for metal ion binding. The same is possible with NH_4^+ as a substitute for monovalent ions[84]. However in both cases exchange with the bulk is so fast that NOEs are observed only to the bulk resonance of the cobalt(III)hexammine or NH_4^+ protons.

Both paramagnetic line broadening by Mn(II) and NOEs to cobalt(III)hexammine have been used as weak distance restraints in molecular dynamics simulations to infer binding sites[81,152].

1.3 Ribozymes – RNA catalysts

When RNA was found to be able to catalyse chemical reactions[7,153] — i.e. to act as an enzyme, a feature previously only ascribed to proteins — the idea of an RNA world was born[154]. Not undisputed, the hypothesis states that in a primordial world RNA was the molecule that both stored the genetic information and performed enzymatic and structural functions — roles that were only later taken over by the more specialized DNA and proteins.

There is evidence that today's ribonucleoproteins (RNPs) in eukaryotic cells might have evolved from pure RNA molecules. RNase P occurs as an RNP in the cell, but the RNA component by itself is catalytically active, classifying it as a true ribozyme[7,155]. Similarly it has been proposed for the two huge protein-RNA complexes, the ribosome and the spliceosome, that the RNA component is more than a structural component and directly involved in the catalytic mechanism[156–163]. Furthermore, the spliceosome shares structural features and the catalytic mechanism with group II intron ribozymes, indicating a common ancestor[107,164–166]. What we learn about ribozymes might therefore enhance our understanding of the evolution and function of important components of the cellular machinery of eukaryotes.

Except for the amide bond formation of the ribosome, all known natural ribozymes catalyse phosphoryl transfer reactions, even if in vitro selection has extended this repertoire[167–172].

In most small self-cleaving ribozymes[121,173], the 2'OH nucleophilically attacks its adjacent phosphate leading to a 2'-3' cyclic phosphate. To this group belong the hairpin[174], hammerhead[175], HDV (Hepatitis Delta Virus)[176] and VS (Varkud Satellite)[177] ribozymes which are ubiquitously employed by viruses to cleave the concatenated genomic RNA. They all catalyse the same reaction, but use different strategies to achieve the necessary alignment of nucleophile, scissile phosphate and leaving group[121,178]. There is evidence for nucleobases acting as acid-base catalysts[179–183] and catalysis does not strictly require divalent ions[179,184–186], even if without them it is less efficient and dependent on high concentrations of monovalent ions[128,187,188]. Three other small self-cleaving ribozymes have more recently been discovered in prokaryotic mRNA (glmS) and eukaryotic pre-mRNA (CPEB3 and CoTC)[189]. The glmS ribozyme in the 5'-UTR of the bacterial glmS gene acts as a riboswitch by cleaving in response to substrate binding[190,191], the HDV-like CPEB3 ribozymes are found in introns of all kinds of higher organisms[192,193] and the CoTC ribozymes in the 3'-UTR are expected to play a role in transcription termination[194].

The large group I and group II intron ribozymes[195–198] and RNase P[199] are strictly dependent on divalent metal ions, which are suspected to directly participate in catalysis via a two-metal-ion mechanism[103,107,123,124,200,201] (see Figure 1.9) mechanistically equivalent to that of protein-based phosphoryl transferases[202]. They catalyse phosphodiester cleavage resulting in 3'OH and 5'-phosphate ends and in the case of the self-splicing reaction of group I and II introns also subsequent ligation of the exons[203,204]. RNase P catalyses the removal of a 5'-terminal leader of tRNA[205]. Apart from the ribosome and some artificial ribozymes it is the only true enzyme, meaning that it performs multiple rounds of catalysis *in trans*. For completeness also the group-I-like cleavage ribozyme has to be mentioned. It is related to group I introns and catalyses the formation of a 2'-5' phosphodiester bond, similar to the first step of splicing, resulting in a 3-nt lariat which can serve as a protecting cap[206,207].

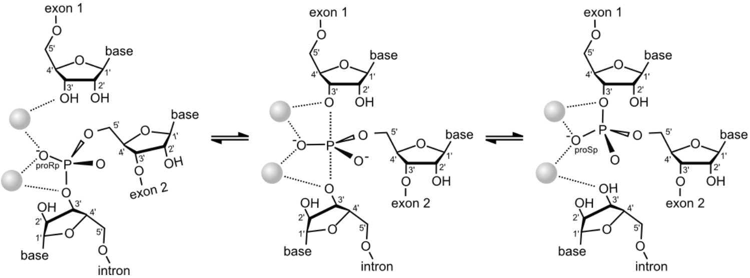


Figure 1.9 The two-metal-ion mechanism of phosphoryl transfer[123], here using the example of the second step of splicing in group II introns (adapted from[208,209]). The two Mg(II) ions activate the nucleophile and stabilize both transition state and leaving group.

1.4 Group II introns

Group II introns are among the largest known ribozymes (1000 and more nt, for reviews see[197,198]). They have first been found in bacteria, archaea and the organelles of lower eukaryotes, but recently also in mammals[210–216]. They catalyse their own excision from an RNA strand in two consecutive phosphotransesterification steps (Figure 1.10a). In the first one a distal 2'OH or H_2O is the nucleophile attacking the phosphoryl at the 5' splice site, resulting in a lariat or linear intron, respectively. In a second step the released 3'OH of the 5'-exon is the nucleophile in a reaction that ultimately joins both exons and releases the intron. Both steps require divalent metal ions and are reversible[164,165,217]. With the help of protein co-factors, reinsertion into RNA and DNA is possible allowing group II introns to participate in the modulation of the genome as mobile genetic elements[218–221]. Insertion into DNA is strictly sequence-specific in some group II introns and thus of potential use in gene therapy[222–226]. A phylogenetic relationship of group II introns with retrotransposons[227] as well as with the eukaryotic spliceosome has been proposed[160,166]. Many group II introns are catalytically active *in vitro* in the presence of Mg(II). Usually, however, significantly higher metal ion concentrations than the physiological ones are required for efficient splicing[228] (ranging from 5 to 100 mM Mg(II) in addition to 0.5 to 2 mM monovalent salt in different introns). *In vivo* most group II introns are stabilized by intron- or host-encoded protein co-factors, which relieve part of the Mg(II) requirement[218,229–232].

Group II introns are divided into the three major classes IIA, IIB, IIC[213,233],

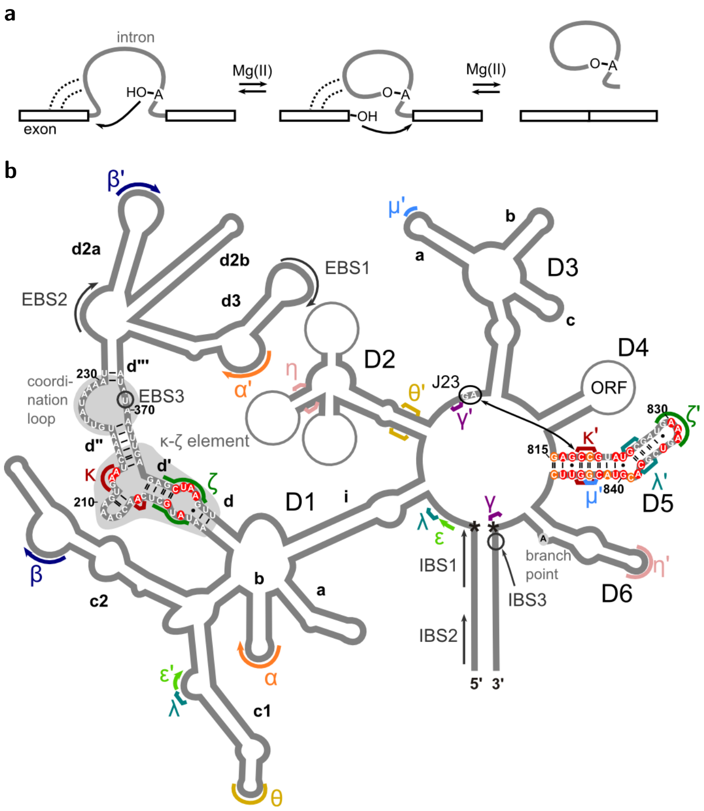


Figure 1.10 Group II intron function and architecture. a) The two steps of splicing performed by group II introns. In a first step the 5' splice site is attacked by the 2'OH group of the branching adenine. In the second step the released 3'OH of the 5' exon attacks the 3' splice site, releasing the intron lariat and joining the exons. b) The secondary structure of Sc.ai5γ, an example of a group IIB intron. The primary sequences are only shown for D5 and the κ-ζ region plus adjacent coordination loop. Nucleotides that are conserved or partly conserved are highlighted in red and orange, respectively. Asterisks indicate the splice sites. Previously identified tertiary contacts are indicated by greek letters and the interaction of J23 nucleotides with the lower helix of D5 is denoted by a black arrow. EBS1,2,3 ... exon binding sites. IBS1,2,3 ... intron binding sites. ORF ... open reading frame.

in addition to several less typical bacterial classes[233–235]. The only available group II intron crystal structures are from a group IIC intron[105,106,236–238]. This class is the most ancient lineage and occurs only in bacteria. Its members are only about half the size of group IIA and IIB members, but they still share major active-site components[233] and so the crystal structures were able to confirm many of the tertiary contacts that were previously predicted on the basis of group IIB introns[239–241] (Figure 1.10b and next section).

Group II introns are characterized by a highly conserved secondary structure consisting of six independently folding domains[213,242] (Figure 1.10b). Domain 1 (D1) is the largest and serves as a scaffold for the assembly of the others[243,244]. It contains the exon binding sites (EBS) which recognize short complementary sequences in the exon to determine the splice site[245]. D2 might play a role in architectural assembly, forming tertiary contacts to D1 and D6[246,247] and D3 is not essential for, but significantly increases, catalytic activity[248–250]. D4 can contain an open reading frame coding for protein co-factors[251]. D5 is the phylogenetically most conserved domain and contains catalytically essential functional groups, while D6 contributes the attacking 2'OH nucleophile for the first step of splicing.

1.4.1 Domain 5 – the central element in group II intron catalysis

While several other domains — excluding D1 — can be deleted without completely abolishing catalytic activity, D5 confers activity even when added *in trans* to the rest of the intron[252–255]. Many important findings have been made using constructs of the group IIB intron ai5γ (Sc.ai5γ, Figure 1.10b) located in the cytochrome oxidase 1 gene of *S. cerevisiae* mitochondria: like D135 (Domains 1, 3 and 5, connected by only truncated helices of D2 and D4), which is catalytically active and has been used in folding and metal cleavage experiments (e.g.[256–258]) or the two-partite system exD123+D5/D56 (the 5'-exon attached to D123 and D5/D56 *in trans*), which proved a valuable tool to characterize the first step of splicing[252,255,259]. Even the minimal combination of D1 and D5 alone is able to catalyse the cleavage of small substrates[252,254,260], whereas the branch site of D6 is not required, because this first step of splicing can also proceed by hydrolysis, leading to a linear intron[252,261,262].

All components important for catalysis are brought together at Domain 5

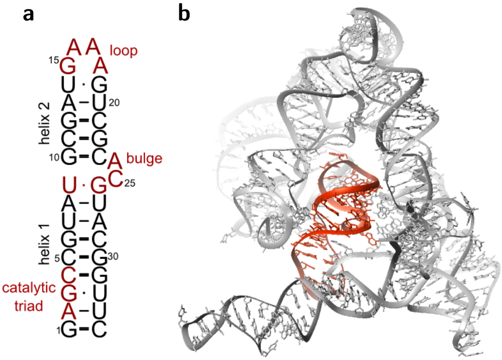


Figure 1.11 a) Domain 5 of the Sc.ai5γ group II intron, indicating important secondary structure elements. b) D5 (in red) is embedded at the core of the *O. iheyensis* group IIC intron crystal structure (PDB-ID: 3G78[237]).

Domain 5 (D5) is the most conserved part of group II introns in sequence and secondary structure. It is typically 34 to 35 nt long, capped by a terminal loop, with two helical regions that are interrupted by a two-nucleotide bulge (Figure 1.11a). Three nucleotides at the base of the lower helix are known as the "catalytic triad"[263]. Especially the practically invariant G (G3 in Figure 1.11a) at its center is crucial for catalysis[213,264], but also the base pairing of the flanking bases[263,265] and especially the phosphates of the first and second nucleotide and the 2'OH of the third nucleotide[266] as well as major groove functionalities[267] are important. Nucleotides at the opposite strand are less critical. While the central base pair is usually a G-U wobble, also G-C and G-A mutations support catalytic activity[263,265].

The two-nucleotide bulge is the second region in D5 that plays an important role in group II intron function. While by far the most common bulge nucleotides are AC, followed by a G engaged in a G-U wobble pair, sequence seems to be not as critical as a minimum of two unpaired nucleotides[268]. In particular phosphates and 2'OH groups, but also nucleobases at and close to the bulge are involved in binding and catalysis[266,269,270].

In the now available group IIC crystal structures[106,236,238] D5 is embedded in the very core (Figure 1.11b). This central position was not unexpected, considering the multiple tertiary contacts that had been established previously (Figure 1.10b). The terminal loop of D5 interacts with a receptor at the center of D1, frequently in a canonical GNRA tetraloop receptor interaction[271] (ζ-ζ', Figure 1.10b). The four nucleotides of κ in D1 also fold into a GNRA motif as predicted by Boudvillain and Pyle[266] and now confirmed in the group IIC crystal structure and interact with the minor groove of tandem G-C pairs in the lower helix of D5[272]. A contact of D3 to the ribose of one of the κ' Gs has been termed μ-μ'[273]. Two G-C pairs in the upper helix of D5 have been shown to interact with nucleotides in the c1 helix of D1 and close to the 5'-splice site[274] (λ-λ', Figure 1.10b) and also the adenine in the D5 bulge has been linked to the 5'-terminal region of the intron[239]. Cross-linking studies[239,275] connected catalytic triad residues to the linker between D2 and D3 (J23), a region which also makes a contact to the 3'-splice site[276] (γ-γ'). Later studies established connections from this linker to the D6 branchpoint[239,240] thus completing the network of interactions that brings together 3'- and 5'-splice site with the branchpoint and the catalytically essential groups of D5[239,277,278].

D5 in the crystal and in solution

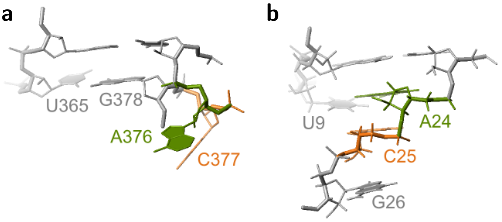


Figure 1.12 The structure of the D5 bulge in the crystal structure of a group IIC intron from *O. iheyensis* (PDB-ID: 3G78[237]) (a) and in D5 from Sc.ai5γ in solution (PDB-ID: 2LPS[279]) (b). While the sequence is the same in both structures, conformations are different. Adenine and cytosine (green and orange) are looped out in the crystal and stacked in solution. The G-U wobble pair observed in crystal structures (U365-G378) is not formed in solution (U9-G26).

The previously established network of tertiary interactions involving D5 is now illustrated in a set of crystal structures of a group IIC intron from the bacterium *O. iheyensis*[105,106,236–238]. The first set of structures[105,236,237] shows the intron in a post-catalytic state with and without the two joined exons. Mutation of the central G in the catalytic triad of D5 then yielded the structure of an inactive intron and a first glimpse at its pre-catalytic state[238]. Most recently, Marcia and Pyle[106] crystallized the intron in various metal ion conditions and were able to contribute structures corresponding to various states during the catalytic reaction.

There is no sign of large-scale conformational changes in any of the structures. They all, however, lack electron density for D6, which has been proposed previously to relocate between the two steps of splicing[246]. The conformation of D5 and the active site does not change markedly in any of the structures.

The two nucleotides in the J23 linker and C377 in the bulge of D5 form the third strand of a triple helix in the major groove of the catalyic triad. At the bulge the lower and upper helices stack co-axially and the phosphate backbone at the solvent-exposed A376 and C377 makes a sharp turn downwards towards the backbone of the catalytic triad residues (Figure 1.12a and 1.13), which was similarly observed already in a previous crystal structure of a Sc.ai5γ D5D6 construct[280]. In the post- and pre-catalytic structures of the group IIC intron the scissile phosphates of the joined exons and of a strongly kinked 5'-splice site, respectively, are positioned right next to this region of concentrated negative charge (Figure 1.13a,b). The adenine in the bulge stacks with G5 at the 5'-end of the intron, supporting the connection to the 5'-splice site. The conservation of the G in the G-U wobble below the bulge (G26-U9 in Figure 1.11a, and G378-U365 in Figure 1.12) is explained by a base-pairing interaction to an adenine in the D1 stem.

Two NMR structures of D5 reveal a striking plasticity of the bulge region[279,281,282] (Figure 1.12). In contrast to the solvent-exposed position in the crystals, the adenine is stacked in, while the guanine, which forms the G-U wobble adjacent to the bulge in the crystal structures, is exposed to the solvent[281,282]. The flexibility in this region[44] might be required to allow the global contortion experienced by D5 upon docking to the rest of the intron[277,283].

D5 harbours the catalytic metal ion binding sites

The backbone of the D5 bulge and catalytic triad in the *O. iheyensis* crystal structures can coordinate two Mg(II) ions[106,236,284]. Consistent with the predicted two-metal-ion mechanism[123,164] they are 4 Å apart and face the scissile phosphate bond. The Mg(II) ions make inner-sphere contacts to non-bridging phosphoryl oxygens of the first two nucleotides of the catalytic triad, the bulged C377 and U375 adjacent to the bulge (Figure 1.13). This is consistent also with previous studies on the Sc.ai5γ intron in solution by phosphorothioate interference, metal ion switch experiments and Tb(III) cleavage[75,256,266,269]. In the post-catalytic state[236], coordination to the O3' of the 5'-exon (possibly also the O2') and to the non-bridging phosphoryl oxygens of the scissile phosphate (Figure 1.13a) reflects well the endpoint of the proposed reaction mechanism[164,165,208,209,285] (compare to Figure 1.9). In the structure representing

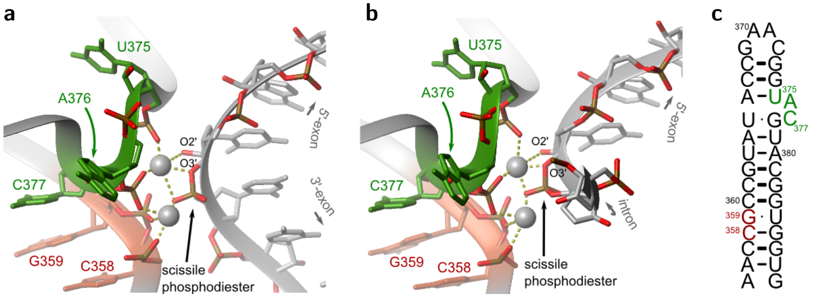


Figure 1.13 Inner-sphere coordination of two Mg(II) ions in the active site of a group IIC intron in its (a) post-catalytic (PDB-ID: 3G78[237]) and (b) pre-catalytic (PDB-iD: 4FAQ[106]) state. The active site remains remarkably unchanged. A two-nucleotide bulge (green) introduces a sharp turn in the backbone of D5 and leads to the close approach of several phosphate groups, creating a favourable binding pocket. c) Secondary structure scheme of OiD5 indicating the catalytic triad (orange) and bulge (green) nucleotides shown in a) and b).

the pre-catalytic state[106] (Figure 1.13b) Ca(II) was used instead of Mg(II) in the crystallization conditions because it is known to inhibit catalysis[67]. The active site geometry does not change and two Ca(II) ions replace the Mg(II) ions observed in the post-catalytic structure. Even if the substrates are different, the scissile phosphate is presented in both pre- and postcatalytic structures in exactly the same way.

Specific Mg(II) interaction at the bulge[281,282] and the catalytic triad[282] was also observed in the isolated D5 in solution, even though the bulge conformation differs significantly from that in the crystal structure (see previous section). Strong effects on the chemical shift of imino protons at these sites are also induced by cobalt(III)hexammine, a mimic for hexahydrated Mg(II)[267]. In solution Mg(II) is also observed to exert a very moderate stabilizing effect on the bulge adenine of the isolated D5 of group II intron LSU/2 from *P. littoralis* (Pl.lsu/2), while the central G of the catalytic triad shows enhanced dynamics with Mg(II)[44]. A slightly more rigid D5 molecule and small conformational changes in the bulge region can be deduced from Raman difference spectroscopy in the same construct[286].

The parallels to spliceosomal snRNA U6

The spliceosome is a megadalton complex consisting of a multitude of proteins and five snRNAs which are expected to catalyse the splicing reaction[287].

Parallels with group II introns in the splicing mechanism[165,288–290] as well as similarities in structural components[291,292] indicate the evolution from a common ancestor, which in the case of the spliceosome involved fragmentation and recruitment of protein co-factors.

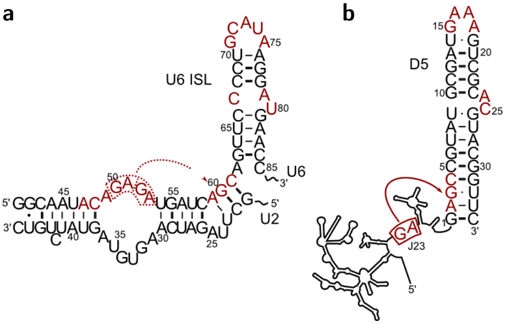


Figure 1.14 Similarities in the yeast U2-U6 snRNA and D5 of group II introns. One of the two GA in the ACAGAGA box (a) could interact with the adjacent AGC triad in a similar way as the GA in the J23 linker region in group II intron Sc.ai5γ (b).

While it is difficult to dissect the spliceosomal components directly involved in catalysis, much points to the complex of snRNAs U2 and U6 as the catalytic core[160,293–296], a structure that shares striking features with group II intron D5[107,166]. The intramolecular stem loop of U6 snRNA (U6 ISL)[292,297] contains a bulged uracil which is flipped out in response to the protonation of the adjacent A^+-C pair with a pK_a of 6.5[297–299] (Figure 1.14). This bulge is separated from a conserved AGC triad by the same number of base pairs as bulge and catalytic triad in D5[263,300,301] (Figure 1.14). The parallels to D5 extend to the coordination of catalytically important metal ions at U6's bulge and AGC triad[164,165,285,288,293,297,302,303] and in an *in vivo* splicing assay U6 can be replaced by D5[159]. A proximal ACAGAGA sequence[292,304] has been proposed to provide the equivalent of the J23 sequence that forms a triple helix with the catalytic triad in group II introns[166]. A U2-U6 construct providing U6 ISL, the AGC triad and the ACAGAGA sequence is sufficient to catalyse a Mg(II) dependent splicing reaction[163,296,305].

1.4.2 Group II intron compaction and the κ-ζ folding control element

The folding of large RNAs is often referred to as "hierarchical", meaning that first the secondary structure is formed, a process which requires only monovalent salt and is usually fast, before tertiary structure falls into place. This second part is usually the focus of ribozyme folding studies. In group II introns it occurs in an apparent two-step mechanism with an early rate-limiting step[257,306]. Folding proceeds slowly (1 min^{-1}) via obligate kinetic intermediates but apparently without kinetic traps to a compact and active structure[257,307,308]. Folding is reversible and unfolding occurs via a single, cooperative transition[306]. The early rate-limiting step was identified to be the folding of D1 because it folds first, independently and with the same Mg(II) dependence as the whole ribozyme[243,307–309].

While most of the folding studies[278,306,308] were performed on the Sc.ai5γ-derived D135 and D1356 ribozymes with optimal *in vitro* conditions of 42°C, 500 mM KCl and 100 mM Mg(II), Fedorova et al.[310] show that the compact intermediate is also reached at more physiological 3 mM Mg(II) and 30°C, albeit at a much slower rate. D3 and D5 are not docked in these conditions, but do so rapidly when more Mg(II) is added[310]. When Mg(II) is diluted out later, the structure unfolds to the compacted intermediate. *In vivo* group II intron folding is facilitated by protein co-factors, which also promote splicing *in vitro* under near-physiological conditions[231,311,312]. The host-encoded DEAD-box protein Mss116p is required for efficient splicing of all group I and group II introns in yeast mitochondria[230,313]. The protein has helicase activity and consumes ATP, but there is still controversy if it stimulates group II intron folding by unwinding misfolded parts or by stabilizing on-pathway intermediates[230,231,312,314–319].

In particular one small region between the two big lobes of D1, the κ-ζ element containing the docking site of D5, and the adjacent coordination loop[239] (Figure 1.10b) was found to play the most important role in the Mg(II)-dependent compaction of D135[258]. In a nucleotide analog interference mapping (NAIM) assay Waldsich and Pyle[258] found that most of the functional groups important for D1 compaction are situated there and designate it a “folding control element” which triggers compaction and folding of group II introns upon its stabilization by Mg(II) or protein co-factors. In a subsequent time-resolved NAIM study, mutations in the κ-ζ element were shown to determine also the rate of compaction[320]. At low Mg(II) concentrations that do not support stable compaction, the adjacent helices (d, d’, d”, Figure 1.10b) were found to be unstable and particularly sensitive to Tb(III) cleavage, which is a sign of the flexibility of the backbone in the region, but also indicates a larger electronegative potential there. The slow rate of compaction despite the absence of kinetic traps could thus be due to an unstable native conformation that needs to be captured by the binding of Mg(II)[320]. An unstable region around the κ-ζ element and the coordination loop was also observed *in vivo* in *S. cerevisiae mss116*-knockout cells[321], where the intron (the full D123456) forms most of its secondary structure but lacks the tertiary fold.

1.5 Thesis outline

This work focuses on the two arguably most central and important parts of group II intron ribozymes, their structure in solution and their interactions with metal ions.

In the first part of the project the solution structure of domain 5 from *A. vinelandii* intron 5 (AvD5, Figure 1.15a) provides the first structural glimpse on an intron from the bacterial class E (Figure 1.15b). A conserved G-A pair sets it apart from previously studied related hairpins (Figure 1.15c). It is remarkable not only because it involves the almost invariant G in the catalytic triad, it is also unusual to find such an isolated G-A mismatch embedded in what is a regular A-form helix in other group II introns. Together with previous structural data on D5 from other organisms and U6 ISL (Figure 1.15c) our findings from NMR and fluorescence experiments can help us understand the dynamic processes involved in the transition from the free to the intron-docked form.

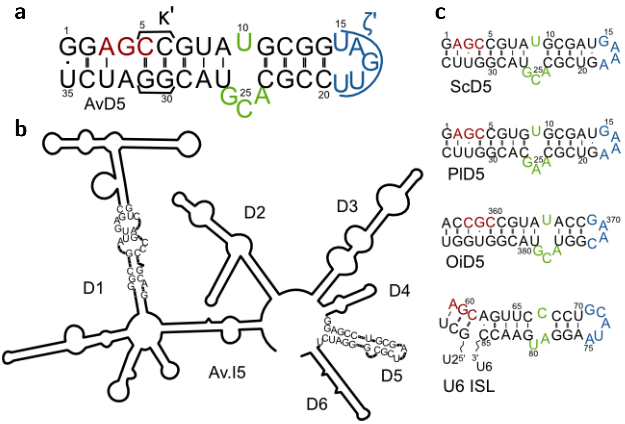


Figure 1.15 a) Secondary structures of the AvD5 construct used in the first part of this thesis. b) Secondary structure of the whole *A. vinelandii* intron 5 (Av.I5)[235]. In D5 and in the region in D1 that is expected to bind D5, nucleotides that are conserved in bacterial class E introns are specified. c) Related D5 hairpins for which structural data exists. ScD5: D5 from Sc.ai5γ[281], PlD5: D5 from Pl.lsu/2[282], OiD5: D5 from an *O. iheyensis* group II intron[236], U6 ISL: stem loop from U2-U6 snRNA complex of the spliceosome[297].

In the second part of the project the κ-ζ element from the more well-studied Sc.ai5γ intron (Figure 1.16) is characterized in solution. This crucial region governs the initial compaction of D1[258,320] to later provide the binding platform for D5[272]. Employing NMR methods we identify the conformational equilibria and the localized interactions with metal ions that will ultimately govern the first steps of folding of the intron.

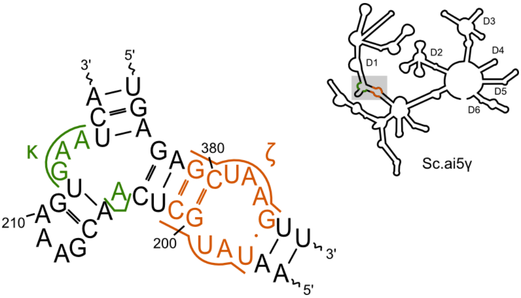


Figure 1.16 The κ-ζ element of the Sc.ai5γ intron was studied in solution in the second part of this thesis. Tertiary interaction motifs κ (green) and ζ (orange) are indicated.

2

Domain 5 of a bacterial group II intron

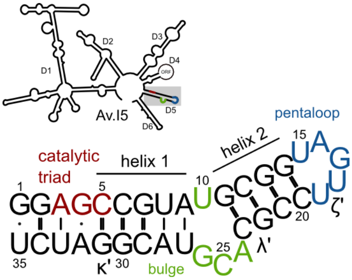


Figure 2.1 Secondary structure of AvD5. The inset indicates its position in the secondary structure of Intron 5 from *A. vinelandii* (Av.I5).

We have studied the full wild-type form of domain 5 from *A. vinelandii* intron 5 (we refer to it as AvD5) corresponding to residues 1836-1870 of the intron . The "catalytic" domain 5 of group II introns, while usually no more than about 35 nucleotides long, contains a multitude of residues indispensable for catalytic funtion of the intron. The AvD5 hairpin is capped by a UAGUU-pentaloop and the Watson-Crick base paired regions are interrupted by an asymmetric internal loop (always referred to as the "bulge" because two possible canonical base pairings of U10 to A24 or G26 would result in a 2-nucleotide bulge) and a G-A mismatch in the region known as the "catalytic triad" (A3 to C5, Figure 2.1). The bulge region is shared with the much more well studied D5 from *S. cerevisiae* group IIB intron ai5γ (ScD5, compare in Figure 1.15c of the Introduction), while the G-A pair in the catalytic triad is a feature specific to the bacterial class E of group II introns. In other introns usually there is a G-U pair at this position. Single G-A mispairs in this sequence context are not found inside standard RNA helices in the protein data bank (PDB)[322], nor has the UAGUU pentaloop been investigated in terms of structure and metal ion binding so far.

2.1 Spectral features and assignment

The correct assignment of proton, carbon and nitrogen resonances in NMR spectra is the basis of all the following structural and metal ion binding studies. It allows to precisely localize the information on distances and dihedral angles to the corresponding parts of the molecule and thus to determine local structure. In addition it makes it possible to record the influences of metal ions and other factors on a per-atom basis.

To find the optimal conditions for NMR experiments, 1D ^{1}H spectra were recorded at several KCl concentrations (0 to 200 mM KCl) and temperatures (290 K to 310 K). Peaks tend to become sharper at higher temperatures, but the effect is not very pronounced (see Appendix 1a). 290 K proved to be the temperature where G26H2' is least likely to be lost in the water signal. In general experiments were performed at 290 or 300 K, unless they involved exchangeable protons, for which lower temperatures are necessary. The NOESY peak pattern does not change significantly in this temperature range. A comparison of chemical shifts proves the unpaired regions — around the G-A mismatch, the bulge and especially the loop — to be most sensitive to temperature changes (Appendix 2a).

Peaks are equally sharp over a wide range of KCl concentrations (Appendix 1b). From the point of view of the imino protons, already the addition of 20-40 mM KCl leads to maximally sharp peaks. Addition of more KCl gradually broadens the peaks, but even at 150-200 mM the effect is not very strong (Appendix 1b). The chemical shifts in the bulge region are the most influenced by KCl concentration (Appendix 2b) and in contrast to the rest of the molecule the non-exchangeable protons in this region broaden noticeably when the KCl concentration is raised from 60 to 120 mM. All further spectra of AvD5 were then usually recorded in the presence of 60 mM KCl.

2.1.1 Sequential walk and assignment of non-exchangeable protons

The assignment of AvD5 is primarily based on ^{1}H,^{1}H-NOESY spectra in D_2O. Most of the H6/H8, H1' and H2' protons can be attributed by following the sequential intra- and intermolecular H6/H8-H1' and H6/H8-H2' peaks in NOESY spectra of partially deuterated samples. At pH 6.7-6.8 (these slightly acidic conditions are commonly used to protect the RNA from auto-cleavage) the sequential walk is interrupted only in the region of the catalytic triad between A3H8-G2H1' and C6H6-H1' and between A32H8-H1' (sometimes already at G31H8-H1') and C34H6-H1' (and likewise for H2' resonances) (Figure 2.2). Many residues in the region of the bulge — from A9 to G11 and C23 to U27 — are broadened. A24H8 and H1' are often hardly visible at all.

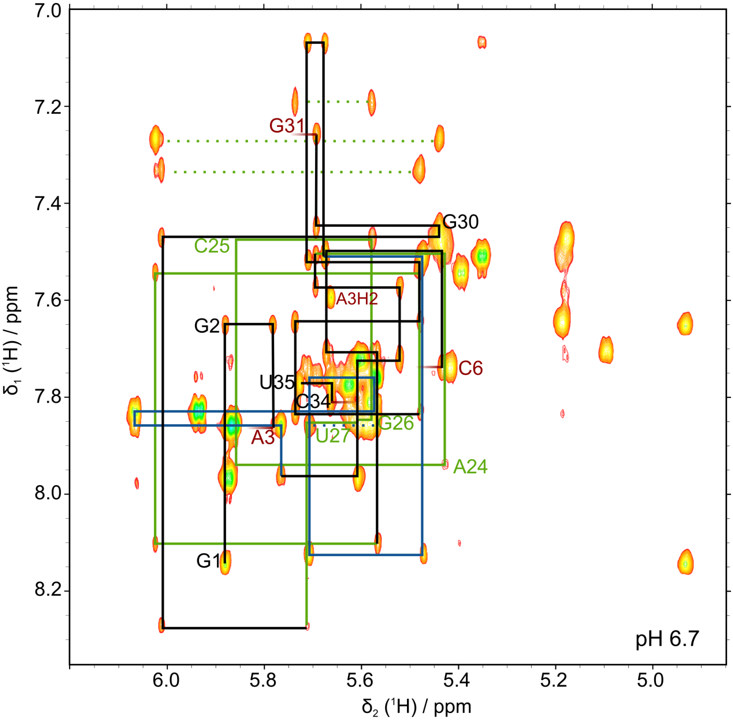


Figure 2.2 Sequential walk region of a ^{1}H,^{1}H-NOESY at pH 6.7. H2 resonances are indicated as interrupted lines. The sequential walk is interrupted between A3 and C6 and between G31 and C34. Resonances in the region of the catalytic triad are labelled in red, in the region of the bulge in green and in the loop in blue as indicated in the secondary structure scheme. The insets show details of the G17H1' and U18H1' doublets. (1 mM AvD5, 300 K, 700 MHz, 100 mM KCl, 10 µM EDTA)

In an attempt to stabilize the end region and potentially exert a favourable influence on the neighbouring unstable catalytic triad region we also recorded spectra of a U35C mutant of AvD5. However, apart from minor variations in the proton chemical shifts of the first two base pairs, the substitution has no effect.

Sequential walk assignments are supported by ^{1}H,^{13}C-HSQC spectra of the aromatic carbons (to distinguish H6, H8 and H2 protons) and ^{1}H,^{1}H-TOCSY spectra (to identify the pyrimidine H6-H5 peaks). Assignment of the remaining sugar protons was supported by ^{1}H,^{13}C-HSQC spectra of the sugar region and by w$_1$,w$_2$-^{13}C,^{15}N-filtered NOESY and TOCSY experiments of AvD5 samples where one of the four nucleotides was ^{13}C,^{15}N-labelled.

2.1.2 Sugar pucker and other dihedral angles

The most striking features of the NOESY spectra are the G17 and U18 cross peaks which display intense peak doublets in their H1' and H2' resonances, indicating a large homonuclear coupling of H1' and H2' (measuring 6.8 ± 0.3 Hz and 6.4 ± 0.2 Hz, respectively, see Figure 2.2), which is not observed for any other residue. In a purely N-type (C3'-endo) sugar pucker $^3J_{H1'H2'}$ is <2 Hz, in a purely S-type (C2'-endo) conformation it is about 10 Hz. Ribose sugar rings are known to interconvert rapidly between N- and S-type conformation, and the fraction of N- and S-type sugar conformation can be estimated from the measured average $^3J_{H1'H2'}$ values using the following equation[52]:

$$^3J^{av}_{H1'H2'} = (1 - pS) \cdot {}^3J^{N}_{H1'H2'} + pS \cdot {}^3J^{S}_{H1'H2'} \qquad (2.1)$$

where pS is the fraction of S-type conformation. In the case of residues G17 and U18 we obtain values around 70%, indicating that their sugar rings are predominantly, but not solely, S-type.

A linear combination of the sugar ^{13}C chemical shifts of a residue can also be used to determine the sugar pucker[323,324] and clearly puts G17 and U18 in the range of C2'-endo, while the other loop nucleotides are in the C3'-endo range. These observations are confirmed in ^{1}H,^{1}H-TOCSY spectra with 50 ms mixing time where only G17 and U18 present exceptionally strong H1'-H2' peaks and well visible H1'-H3' cross peaks, clearly indicating that their sugars are predominantly S-type. Medium intensity H1'-H2' peaks in the TOCSY spectra indicate that also the sugar puckers of residues G1, A16, U19, C25, G26 and U35, situated at the hairpin ends, in the loop and in the bulge regions, deviate partly from the N-type conformation typically observed in A-form RNA helices (Figure 2.3a).

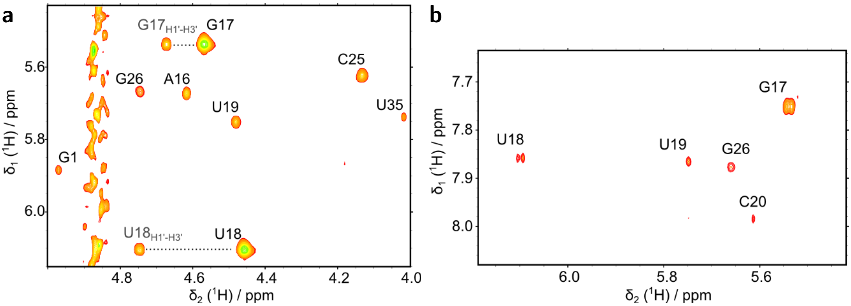


Figure 2.3 Experimental evidence for sugar pucker types and glycosidic angles in AvD5 deviating from standard A-form helix values. a) ^{1}H,^{1}H-TOCSY sugar region at pD 6.7 displaying H1'-H2' peaks for residues with partly S-type sugars. G17 and U18 in the loop are predominantly S-type, which allows to observe also their H1'-H3' cross peaks (50 ms mixing time, 1 mM AvD5, 290 K, 700 MHz, 100 mM KCl, 10 µM EDTA). b) Section of a ^{1}H,^{1}H-NOESY spectrum at 60 ms mixing time, indicating the most intense intraresidue H6/H8-H1' peaks (0.5 mM AvD5, 290 K, 60 mM KCl, 10 µM EDTA, 600 MHz).

In Watson-Crick base paired helical regions nucleotides adopt an *anti*-conformation where the glycosidic angle χ is around -160° [52]. In loops and bulges or non-canonical base pairs *syn*-conformation ($\chi = 60^{\circ}$) can occur, which is characterized by a particularly short intranucleotide H6/H8 to H1' distance, similar to the distance between pyrimidine H5 and H6 protons. In a 60-ms-mixing-time NOESY spectrum, G17 in the loop displays the strongest H8-H1' cross peak, corresponding to a distance of 2.5 Å and indicative of predominantly *syn*-conformation. The U18H6-H1' and G26H8-H1' cross peaks are also relatively strong, yielding distances of 2.88 and 3 Å, respectively, suggesting a glycosidic angle that deviates from the standard *anti*-conformation, but is also not completely or not all the time in *syn*-conformation (Figure 2.3b). U19 and C20 H6-H1' peaks are even weaker, but still have above average intensities. The fact that all those residues (G17, U18, U19, C20 and G26) have C6/C8 chemical shifts that are 1-2 ppm downfield shifted from the usual residue specific ranges is also indicative of *syn*-conformation[325]. All other residues have H6/H8-H1' NOE cross peaks that give distances of about 3.5 Å consistent with *anti*-conformation.

None of the ^{31}P resonances fall out of the normal region, which suggests that no α or ζ dihedral angles are in the unusual *trans*-range.

2.1.3 Fast dynamics in the loop

Increased linewidths in NMR spectra are caused by fluctuations of local magnetic environments at rates intermediate on the chemical shift time scale and thus indicate regions with a certain flexibility[140,144]. Such broadening is observed in AvD5 in the bulge and more severely in the catalytic triad region. It is recognized easily in the spectra and will lead to a more cautious evaluation of the data. Inversely, however, it can not be inferred that exceptionally sharp peaks — as observed in the pentaloop of AvD5 — indicate exceptionally rigid regions, because dynamics happening fast with respect to the chemical shift time scale can lead to very sharp peaks that appear at a population-averaged chemical shift and with averaged intensity[140] (Figure 1.7 in the Introduction). Unrecognized this can lead to NOE distance restraints that correspond to an average structure that might not even be sampled by the molecule at all.

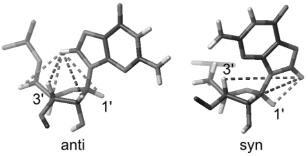


Figure 2.4 Intranucleotide base-to-sugar distances in a guanine in *anti*- and *syn*-conformation.

In the case of G17 at the top of the pentaloop, exceptionally strong intranucleotide base-to-sugar signals are observed (Table 2.1). It is impossible for one single conformation to satisfy both the short H6/H8-H1' and H6/H8-H3' distances at the same time, which strongly indicates conformational exchange. The observed peaks would thus be averages of two or more conformations. Due to the r^{-6} dependence of the NOE on the interproton distance very short distances will dominate the averaged peak intensities, contributing more to the observed peak intensities than their population would suggest[326]. The distances H6/H8-H1' and H6/H8-H3' are much more weakly dependent on the sugar pucker than on the glycosidic angle χ and thus averaging of the sugar conformation is not likely to cause the observed inconsistencies. In contrast, a flipping of the G17 base from *syn*- to *anti*-conformation would result in a very short intranucleotide H8 to H1' distance in one case and short H8 to H3' proton distance in the other (Figure 2.4). The energy barrier which has to be overcome for such an interconversion is especially low for purines with S-type sugar conformation[327].The exceptionally small apparent H6/H8-H2' distances suggest that

Table 2.1 Typical base-to-sugar distances (in Å) in nucleotides with average S- or N-type sugars, respectively, and *syn*- or *anti*-conformation of the base, respectively[52] compared to the distances observed in G17 and U18 of AvD5. The latter distances were obtained from the integration of peaks in 120 ms mixing time ^{1}H,^{1}H-NOESY spectra.

	syn		*anti*			
	S-type	N-type	S-type	N-type	G17	U18
d(H6/H8,H1')	2.5	2.5	3.8	3.8	2.50	2.88
d(H6/H8,H2')	3.9	4.3	3.9	3.3	2.50	1.93
d(H6/H8,H3')	5.1	5.9	3.1	4.9	3.18	2.54

the interconversion takes place across $\chi=0°$, so that H6/H8 would pass close to H2' in the process. This is the less favourable pathway according to Foloppe et al.[327], indicating that conversion in the other direction is sterically hindered. Sugar-to-base distances indicate a similar scenario for U18, but the slightly longer apparent H6-H1' distance indicates a less populated *syn*-conformation.

2.1.4 Assignment of exchangeable protons and base pairing

In ^{1}H,^{1}H-NOESY spectra recorded in H_2O at 275 to 280 K we observe imino signals with cross peaks to the residue on the opposite strand indicating their involvement in stable base pairs for 10 of the 12 expected base pairs (G1-U35, G2-C34, C5-G31, C6-G30, G7-C29, U8-A28, G11-C23, C12-G22, G13-C21, G14-C20) (Figure 2.5). Not observed are imino protons of the expected W-C pairs A3-U33 in the catalytic triad and A9-U27 next to the bulge. We also do not find any evidence for a G4-A32 base pair in the catalytic triad in the ^{1}H,^{1}H-NOESY spectra, even if G4H1 can be tentatively assigned (see below) and none of the bulge residues has observable imino protons. On the other hand, U15H3 can be identified by cross peaks to G14, A16 and U19. It is the only imino resonance in the loop which is observable apart from the tentatively assigned G17 (see below). The first base pair (G1-U35) is always represented by more than one G-U resonance pair due to 3'-end heterogeneity of the RNA. The T7 polymerase employed in *in vitro* transcription tends to attach one or two additional nucleotides at the 3'-end[11,12]. This problem can be greatly alleviated by using methoxy-modified DNA templates[328,329], which we employed routinely in later stages of the project. These overhanging nucleotides, however, also have one advantage: they have a considerable stabilizing effect on the first base pair[37]. Only the faintest imino signals are observable for G1 and U35 in their absence (Figure 2.5 inset).

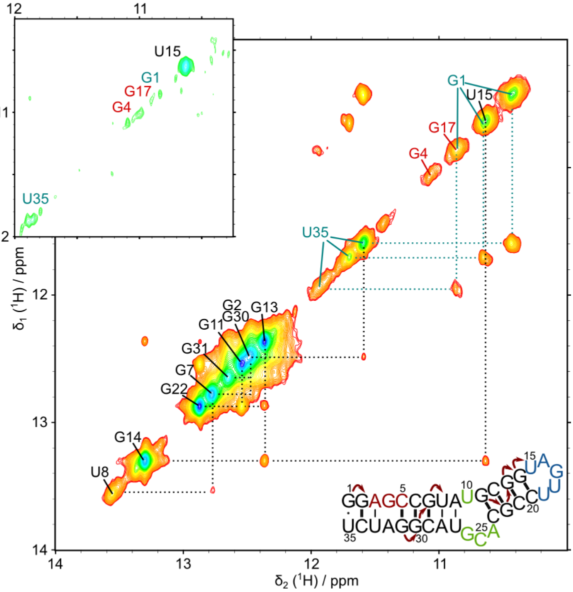


Figure 2.5 Assignment of imino protons in a ^{1}H,^{1}H-NOESY in H_2O at pH 6.7. G4H1 and G17H1 are tentatively assigned in red (see text). Sequential connections that can be made through the imino protons are marked by red arrows in the secondary structure scheme. Multiple resonances for G1H1 and U35H3 due to aberrant 36-nt side products of transcription are marked in cyan. The small inset shows an equivalent spectrum of a sample where the pure 35-nt construct could be separated on the gel. (0.7 and 0.2 mM AvD5 in the big spectrum and the inset, respectively, 280 K, 60 mM KCl, 10 µM EDTA, 600 MHz)

Apart from the 10 base paired imino protons and U15H3, we observe two additional imino resonances at about 10.9 and 11.1 ppm in the region typical for unpaired or non-canonically paired residues. Both of them are also observable in samples with a G1-*C*35 base pair and can thus be distinguished from additional G1 and U35 resonances due to 3'-end heterogeneity. The ^{15}N chemical shifts of the corresponding nitrogens identify them as guanines. Apart from distinct exchange peaks with water, both resonances have only one cross peak each at about 6.4 ppm, a region where an averaged peak of the two protons of a guanine amino group could be expected if the rotation about the C-N bond leads to a sufficiently fast exchange between the two. Only three unassigned guanines remain: G4 in the catalytic triad, G17 in the loop and G26 in the bulge. We tentatively assigned the two resonances to the imino protons of G4 and G17, because G26 can be attributed to another resonance in spectra at higher pH (see Section 2.2.4). Assignment of G4H1 is further supported by a cross peak to A3H8 which can be found at low pH (see Section 2.2.4).

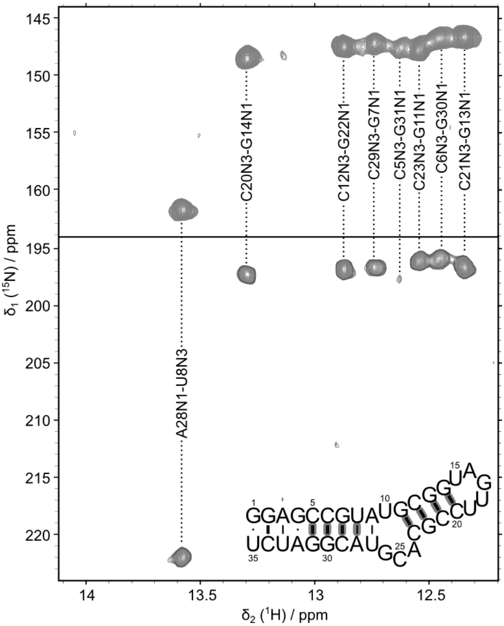


Figure 2.6 ${}^2J_{NN}$ HNN COSY indicating stable W-C base pairs in AvD5 (0.6 mM AvD5, pH 7.8, 60 mM KCl, 10 µM EDTA, 275 K). The confirmed base pairs are marked in the secondary structure inset.

The ${}^2J_{NN}$ HNN COSY can confirm 8 of the 10 base pairs indicated in the NOESY spectra (C5-G31, C6-G30, G7-C29, U8-A28, G11-C23, C12-G22, G13-C21, G14-C20) (Figure 2.6). The G1-U35 wobble pair is missing, because it does not have a N-H···N bond which could be detected by this experiment. The fact that G2-C34 could not be verified either can be due to an overlap with C6-G30. In addition, signals for the second base pair are expected to be weakened anyways because they are affected by the 3' end heterogeneity, which leads to a distribution of the total signal intensity among the several different species. In samples without 3' end heterogeneity on the other hand, the first base pair (G1-U35) is not stable and so the enhanced exchange with the solvent for the second base pair leads to loss of the signal. Taken together, it is not surprising that there is no clear signal for this base pair in the ${}^2J_{NN}$ HNN COSY. However, the observed C34-amino to G2-imino proton NOE cross peaks are an indication that the base pair is formed to a good extent nonetheless.

2.2 Marked pH dependence of catalytic triad and bulge resonances

As described in Section 2.1.1, at pH 6.7 the sequential walk region of AvD5 is interrupted at the catalytic triad G-A mismatch and peaks from the region around the bulge are broader than those from other parts of AvD5. A closer investigation revealed a marked pH dependence in both of these regions.

It turns out that the missing peaks in the catalytic triad between A3 and C5 and between G31 and C34 become observable both at basic and at acidic pH, some of them at very much different positions (see Figure 2.7 and Appendix 3 to Appendix 7). At high pH ($>$7-7.2) both U33H6 and H1' are in unusually shielded positions (at 7.27 ppm and 5.28 ppm, respectively). All A3 and G4 resonances and C5H5 and H6 are comparatively weak as are also G31H2' and A32 and U33 resonances, a fact that does not change even at the highest tested pH of 9.4 and which indicates that there are dynamics that broaden the signals. In many instances we observe a doubling or tripling of peaks in this region, suggesting the co-existence of multiple slowly inter-converting conformations at the G-A base pair. At low pH ($<$5.5-5), U33H6 and H1' have reappeared at 7.79 ppm and 5.61 ppm, respectively (Figure 2.7). G31H2' also has shifted strongly from 4.31 ppm at pD 7.8 to 4.62 ppm at pD 5. Other resonances in the region of the catalytic triad (G4H1', C5H1', C5H6 and G30H8) have shifted downfield less dramatically by 0.1-0.2 ppm (Figure 2.8).

Peaks in the region of the bulge become much sharper above pH 7. However, they disappear completely below pH 6.5 to 6.0 and do not reappear any more like the catalytic triad peaks. Chemical shift changes in the bulge between pH 6.0 and 7.2 are especially marked for A9 and U10, but also G11, A24 and C25 resonances shift (Figure 2.8).

2.2.1 Two pK_a values close to neutral pH

The strong chemical shift changes we observe around pH 7 in the catalytic triad and the bulge region indicate the existence of one or more acid-base equilibria in the neutral pH range. To determine the exact pK_a values we plotted ^{1}H and ^{13}C chemical shifts measured in ^{1}H,^{1}H-NOESY and ^{1}H,^{13}C-HSQC spectra between pH 5.5 and pH 9. Even though many of the curves are compromised — in the bulge, because peaks are not observable at lower pH and in the catalytic

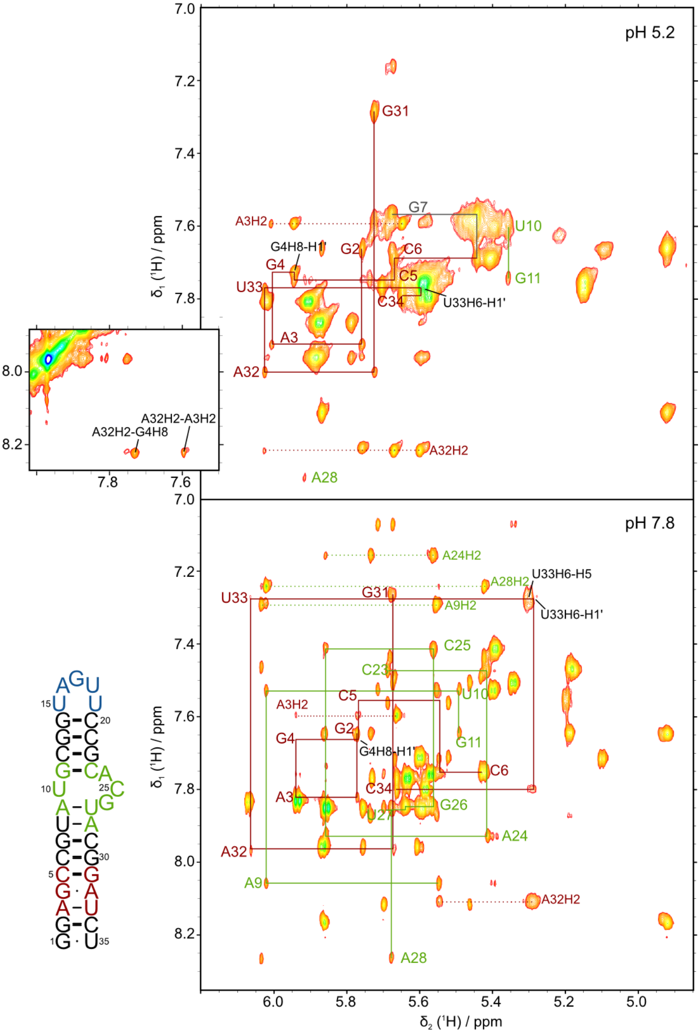


Figure 2.7 ^{1}H,^{1}H-NOESY sequential walk region of AvD5 at pH 5.2 and pH 7.8. Connectivities in the region of the catalytic triad are marked in red, in the region of the bulge in green. Some individual peaks that are mentioned in the text are also indicated in black. Adenine H2 resonances are marked by interrupted lines. In addition, at pH 5.2 G7 is assigned in grey to point out the remarkably sharp peaks in a spectrum where most other resonances from helical regions are broadened due to the beginning acid denaturation. (0.5 mM AvD5, 300 K, 60 mM KCl, 10 µM EDTA, 700 MHz)

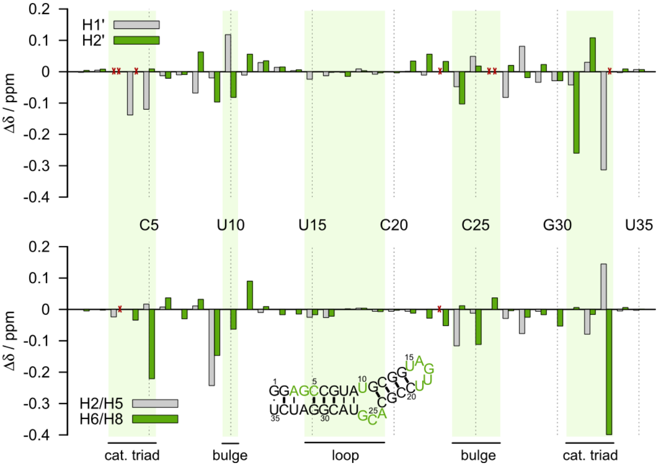


Figure 2.8 Chemical shift changes in AvD5 protons when raising the pH from 6 to 7.2. G31H2' and U33H5/H6 are not yet visible at pH 6, in this case chemical shifts at pH 5 were used instead. To distinguish resonances that are not moving at all from values that are missing because assignment is not possible under all conditions the latter were crossed out red in the graph.

triad, because they are not observable in the intermediate region where the strongest changes happen — reasonably good fits to a Henderson-Hasselbalch derived equation (Equation 4.5) were obtained for a number of proton and carbon resonances (see Table 2.2 for all obtained pK_as and Appendix 9 for the curves in the catalytic triad and bulge regions). The derived pK_a values for the individual protons in the catalytic triad and bulge region, respectively, agree well among each other, i.e. all values in the region A3-C5 and G31-U33 and likewise all values in the region U8-G11 and C23-A28 are similar (see red and green columns in Table 2.2). This is consistent with the assumption of one protonated position within each of the two regions. A protonation event is sensed not only by the concerned residue but also by the resonances of the surrounding nucleotides, all of which can be used to determine the pK_a of the protonation. Taking the mean over all individually determined pK_a values in the catalytic triad and the bulge region (all values marked in red and green in Table 2.2), respectively, we find a pK_a of 6.69 ± 0.10 in the former and of 6.38 ± 0.12 in the latter.

Between the two regions (at C6, G7, C29, G30), not only the chemical shift changes are weak (Figure 2.8) but there is also an overlap of the influences from both sides and so pK_as from there were not included in the average. At the helix ends individual fits indicate an additional higher pK_a (~7.2 which is in the range were the phosphate group of a nucleoside di- or triphosphate is protonated[330]), however, the overlap with influences from the catalytic triad region prohibits accurate quantification. In the pentaloop on the other hand, chemical shifts change pronouncedly only below pH 6. The least-squares fits indicate a pK_a around pH 5, but this can not be determined precisely since we did not go below pH 5 in our titrations.

2.2.2 Where are the protonation sites?

In the previous section we established the existence of two perturbed pK_a values in the regions of catalytic triad and bulge without yet knowing which specific sites are being protonated. To identify them we looked at chemical shift changes of the ^{13}C and ^{15}N resonances in the pH range of the determined pK_as. A protonation event has a significant impact on the electronic environment of the concerned nucleus and will therefore strongly influence its chemical shift and that of its neighbouring nuclei[331–334].

In the catalytic triad region, protonation at A32N1 is demonstrated readily by a -65.2 ppm shift of A32N1 and a -7.5 ppm shift of A32C2 when going from pH 7 to pH 6 (Figure 2.9), agreeing very well with reported literature values: a 70 ppm upfield shift is expected for adenine N1 upon protonation, together with a ~8 ppm upfield shift of the adjacent C2[331,334]. Also among the adenine N3 resonances A32N3 is the one that experiences the strongest shift (>6 ppm downfield, Figure 2.9c).

In the bulge, ^{13}C and ^{15}N shifts upon change of pH indicate protonation at A9N1 (Figure 2.9). In the crystal structures available of D5[236,280], A9 is involved in a W-C base pair and also in the solution structure of ScD5 signs of the formation of the base pair were found[281]. It therefore comes as a surprise that this base and not the unpaired bulge nucleotides A24 or C25 should be protonated, even if we never observe signs of a stable A9-U27 base pair in AvD5. Nonetheless, neither A24C2 nor C25C6 (cytosine C6 are expected to move by ~4 ppm upon protonation of the N3 position[335]) move significantly when the pH is decreased from 7 to 6, while A9C2 and A9C8 clearly do (Figure 2.9a,b). In

Table 2.2 pK_a values from fitting chemical shift changes in ^{1}H,^{1}H-NOESY (a), and ^{1}H,^{13}C-HSQC spectra (b) vs. pH changes in the range of pH 5.5 to pH 9. The catalytic triad and bulge residues that were used to derive the average pK_as for the two regions are marked in red and green, respectively.

a) NOESY

	G1	**G2**	**A3**	**G4**	**C5**	**C6**	**G7**	**U8**	**A9**	**U10**			**G11**	**C12**	**G13**	**G14**	**U15**	**A16**	**G17**
H1'	**6.26**	n.a.[a]	n.a.[b]	**6.66**	**6.67**	**6.70**	6.31	**6.25**	**6.47**[c]	**6.39**			**6.39**	**6.21**	**6.19**	n.a.[a]	**4.89**	**5.10**	n.a.[a]
H2'	**6.97**	n.a.[a]	n.a.[b]	n.a.[b]	**6.94**	**6.74**	n.a.[a]	**6.27**	**6.40**	**6.37**			**6.43**	**6.35**	**6.14**	n.a.[a]	n.a.[a]	n.a.[a]	**5.01**
H6/H8	**7.32**	n.a.[a]	**6.68**	n.a.[a]	**6.66**	**6.6**	6.15	**6.45**	**6.31**	**6.39**			**6.46**	n.a.[a]	**5.3**	**3.01**	**4.69**	**4.87**	n.a.[a]
H2/H5			n.a.[a]		**6.64**	**6.25**		**6.4**[a,c]	n.a.[b]	n.a.[a]				**5.71**			**5.08**	**5.22**	

	U35	**C34**	**U33**	**A32**	**G31**	**G30**	**C29**	**A28**	**U27**	**G26**	**C25**	**A24**	**C23**	**G22**	**C21**	**C20**	**U19**	**U18**	
H1'	**6.62**	**6.46**	**6.60**	**6.65**	**6.65**	**6.52**	**6.34**	**6.39**	n.a.[b]	n.a.[a,b]	n.a.[a,b]	n.a.[b]	n.a.[a,b]	**6.33**	n.a.[a]	n.a.[a]	**4.79**	**5.53**	
H2'	**6.38**	**6.54**	**6.71**	n.a.[b]	n.a.[b]	**6.55**	**6.73**	**6.38**	**6.24**	n.a.[d]	n.a.[a,b]	**6.32**	n.a.[a,b]	**6.40**	**6.22**	**6.07**	**5.42**	n.a.[a]	
H6/H8	n.a.[a]	**6.57**	**6.51**	**6.59**	n.a.[a]	**6.44**[c]	**6.32**	**6.49**	n.a.[a,b]	n.a.[a,b]	**6.29**	n.a.[a,b]	**6.31**	**6.24**	**5.38**	**5.42**	**5.43**	n.a.[a]	
H2/H5	**6.62**	n.a.[a]	**6.63**	**6.75**			n.a.[b,e]	**6.38**	n.a.[b]		n.a.[b]	**6.37**	n.a.[a,b]		**5.67**	**5.22**	**4.63**	n.a.[a]	

b) HSQC

	G1	**G2**	**A3**	**G4**	**C5**	**C6**	**G7**	**U8**	**A9**	**U10**			**G11**	**C12**	**G13**	**G14**	**U15**	**A16**	**G17**
C6/C8	**7.26**	**7.13**	n.a.[a]	n.a.[d]	n.a.[d]	**6.40**	**6.43**	**6.29**	**6.24**	**6.36**			**6.33**	**6.47**	n.a.[a]	**3.55**	**5.31**	**5.09**	**5.24**
H6/H8	**7.30**	n.a.[e]	**6.69**	n.a.[d]	n.a.[d]	**6.60**	**6.05**	**6.54**	**6.44**	**6.30**			**6.39**	**5.97**	**5.86**	**3.65**	**5.43**	**4.94**	n.a.[a]
C2			n.a.[a]						**6.43**									**5.09**	
H2			n.a.[a]						**6.44**									**4.85**	

	U35	**C34**	**U33**	**A32**	**G31**	**G30**	**C29**	**A28**	**U27**	**G26**	**C25**	**A24**	**C23**	**G22**	**C21**	**C20**	**U19**	**U18**	
C6/C8	n.a.[a]	n.a.[a]	n.a.[b]	**6.83**	n.a.[a]	**6.61**	**6.4**	**6.24**	**6.34**	**6.90**	**6.32**	n.a.[a]	n.a.[d]	**6.37**	n.a.[a]	**5.22**	n.a.[a]	n.a.[a]	
H6/H8	n.a.[a]	n.a.[a]	n.a.[b]	n.a.[b]	n.a.[a]	**6.34**	**6.59**	**6.41**	n.a.[a]	**6.43**[c]	**6.27**	n.a.[a,b]	n.a.[d]	**5.84**	**5.44**	**5.62**	**5.76**	n.a.[a]	
C2				n.a.				**6.5**[c]				n.a.[b]							
H2				**6.79**				**6.19**				n.a.[b]							

a … <0.01 ppm ^{1}H or <0.1 ppm ^{13}C chemical shift change
b … two few observable points
c … pH 5.5 titration point was excluded
d … not possible to assign
e … bad fit due to a noticeable second pK_a

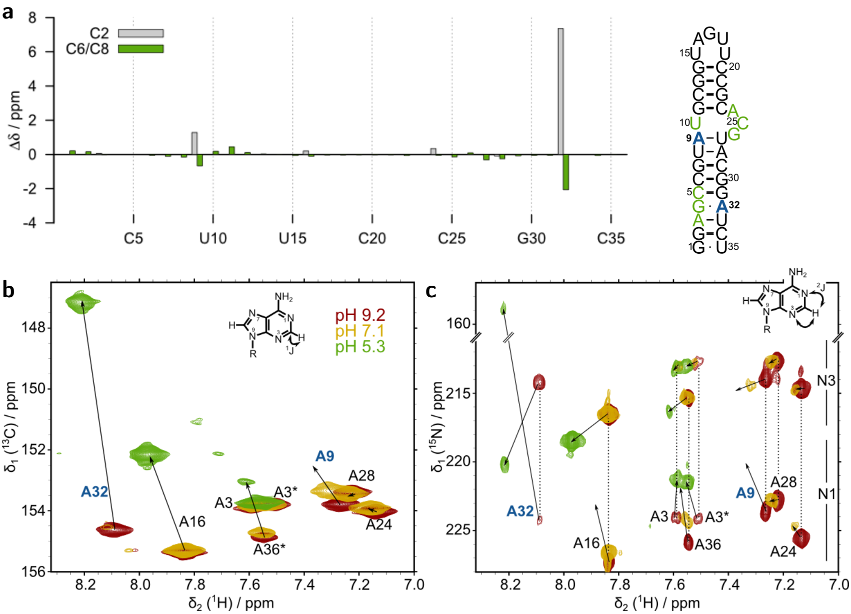


Figure 2.9 a) Chemical shift changes of AvD5 aromatic carbons between pH 7.5 and pH 6.5. b) C2-H2 region in ^{1}H,^{13}C-HSQC and c) N1/N3-H2 region in ^{2}J-^{1}H,^{15}N-HSQC spectra in acidic (green), neutral (yellow) and basic (red) conditions. A9 and A32 are marked in blue in the secondary structure scheme on top (where also bulge and catalytic triad are indicated in green) and their resonances, which experience the strongest shifts, are labelled in blue in the spectra. Assignments for all other adenines in AvD5 are also indicated. Some of them, like A16 and A36 also experience strong shifts, but only at lower pHs. The * marks the duplicated A3H2 resonances due to 3' end heterogeneity. (300 K, 60 mM KCl, 10 µM EDTA)

addition, the fact that A9N1 is the only N1 apart from A32N1 which disappears already at neutral pH — even if it does not reappear at acidic pH like A32N1 — indicates its protonation (Figure 2.9c). This is supported also by A9N3, which after A32N3 is the N3 resonance that is shifting most strongly between pH 9 and 7. The C2 resonances of A16 in the loop and of the 3'-overhang A36, for instance, also move strongly and their N1 disappear (Figure 2.9a,b), but this happens at much lower pH than for A9.

2.2.3 Conformational changes in the bulge

The increasing broadening of resonances in the bulge region is indicative of conformational exchange processes at low pH, even if the complete disappearance of the peaks makes it difficult to assess the nature of the undergone conformational changes. Exchange in the range of milliseconds is indicated by a peculiar cross peak which appears at 13.2 ppm at low pH, involving G11H1 adjacent to the bulge (marked by a green box in Figure 2.10a). The peak does not connect to another observable diagonal peak. While otherwise the missing U10H3 would be the only probable attribution of this resonance, a missing diagonal peak can indicate an exchange cross peak of G11H1 rather than an NOE cross peak. However, since no other cross peaks are present and the resonance can not be observed in ^{1}H,^{15}N-HSQCs, unequivocal identification is not possible so far.

At high pH, on the other hand, a newly formed G-U base pair is indicated by a pair of imino protons (marked in green in Figure 2.10b), which could be assigned to a weakly formed U10-G26 pair. However, due to a lack of additional cross peaks this remains speculative and what exactly happens at the bulge as the pH changes is not deducible from our NMR studies.

2.2.4 Stabilization of base pairs in the catalytic triad region at low pH

At high pH the exchange of the imino protons with H_2O is enhanced and signals for the more weakly hydrogen-bonded protons are lost. In Figure 2.11, this is apparent from the almost complete disappearance of all peaks in the region below 12 ppm, where G-U base pairs, but also in general non-Watson-Crick bonded imino protons are usually found. On the other hand, imino proton resonances from stable W-C base pairs are particularly sharp at high pH.

At low pH we see the appearance of new peaks (Figure 2.10a and 2.11). According to the literature[51,336] the H1 resonances of protonated adenines are expected to be particularly downfield shifted. However, the most downfield shifted peak at 14.4 ppm which appears at acidic pH can be assigned unequivocally to U33H3 (labelled red in Figure 2.10a) by strong cross peaks to G2H1 and A3H2 in ^{1}H,^{1}H-NOESY spectra. A32H1 on the other hand would be expected to have a very strong cross peak to A32H2, but none of the observable imino resonances shows such a correlation. This indicates that the A32 imino proton

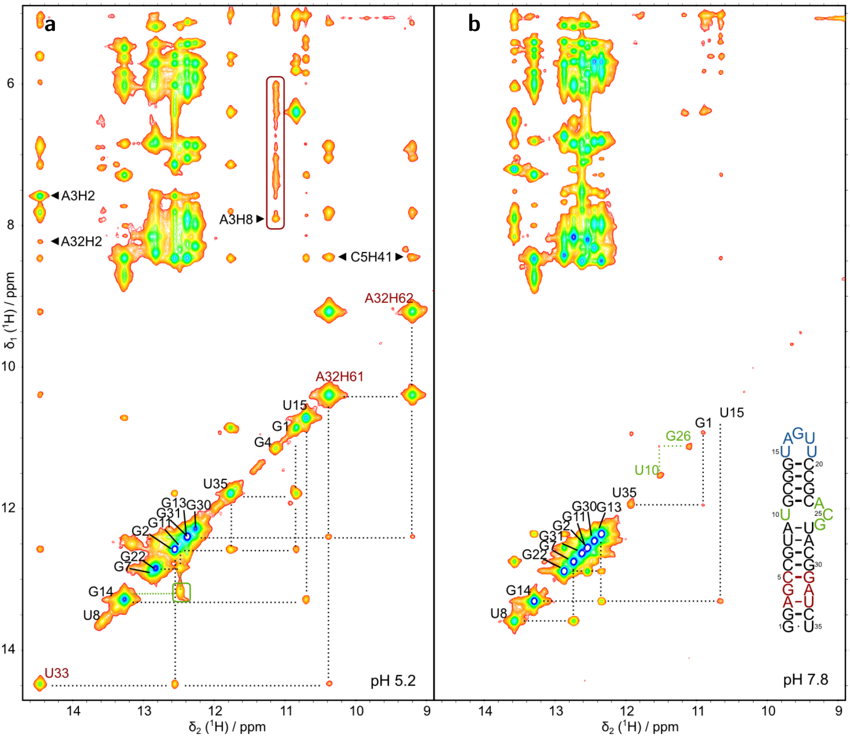


Figure 2.10 Assignment of the exchangeable protons of AvD5 in a ^{1}H,^{1}H-NOESY in H_2O at pH 5.2 (a) and pH 7.8 (b). Imino resonances which are only observable at low or at high pH are labelled in red if they are from the catalytic triad region and in green if they are from the region of the bulge. Unless specified otherwise, the assignments indicate imino protons. a) At pH 5.2 the green box marks a putative exchange cross peak of G11H1. Black arrows and resonance labels in the top part of the spectrum indicate resonances that were important for the identification of the exchangeable proton resonances G4H1, A32H21/H22 and U33H3. The red box indicates a line of cross peaks to G4H1 which are only observable at low pH. b) At pH 7.8 a newly appearing G-U pair (marked in green) is tentatively assigned to G26-U10. In the secondary structure scheme catalytic triad (red), bulge (green) and loop (blue) are indicated. (0.5 mM AvD5, 275 K, 60 mM KCl, 10 µM EDTA)

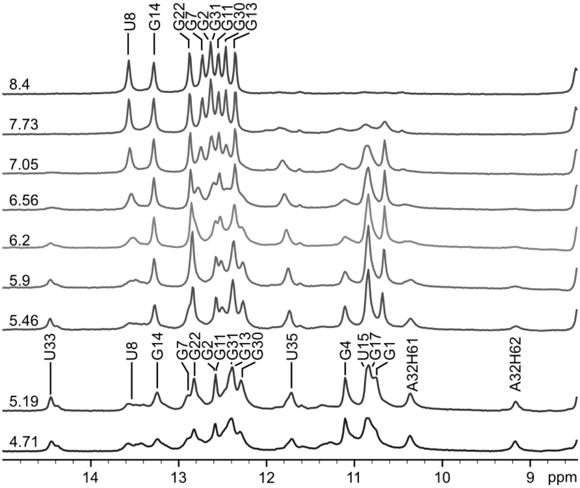


Figure 2.11 1D ^{1}H spectra of AvD5 imino protons as a function of pH. The shoulder observed for U33H3 is due to the 3' end heterogeneity (see Section 2.1.4). All assigned resonances correspond to imino protons except A32H61 and H62. (0.3 mM AvD5, 275 K, 60 mM KCl, 10 µM EDTA)

is exchanging with the solvent water on an intermediate time scale and is therefore not observable. The appearance of U33H3 on the other hand indicates the stabilization of the A3-U33 base pair at low pH.

The chemical shift values of two resonances at 9.1 and 10.3 ppm (labelled red in Figure 2.10a) and the strong cross peak connecting the two indicate a pair of amino protons, which is confirmed by the chemical shift of the attached nitrogen in a ^{1}H,^{15}N-HSQC. Additional cross peaks to both G31H1 and U33H3 assign the amino protons to the G4-A32 mispair. However, the amino protons of G4 would be expected to have a strong cross peak to the G4H1 resonance and such a peak is completely absent. We can therefore attribute the two resonances to A32H61/H62.

The stabilization close to G4/A32 observed through the exchangeable protons at low pH is also apparent in the non-exchangeable ones. Even though some remaining broadness in residues A3, G4, C5 and G31 still indicates a dynamic region, the peaks are better observable at low pH and especially A32 and U33 resonances gain in intensity. Furthermore, resonances belonging to C6 and G7 adjacent to the catalytic triad remain comparatively strong and sharp at pH 5 when a general destabilization of other helical regions is already apparent in an overall broadening and strong downfield shifts of most resonances (Figure 2.7 top).

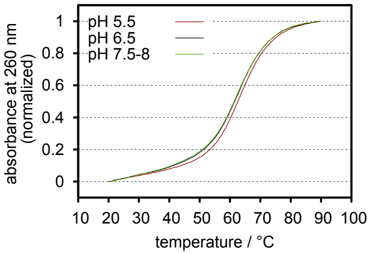


Figure 2.12 UV melting curves of AvD5 at pH 5.5, pH 6.5 and pH 7.5. Each curve represents the average of three independent measurements. Each measurement consisted of 3-4 consecutive up- and downward ramps of the temperature, which were averaged. The absorbance at 260 nm was normalized with respect to initial and final intensities. (1 µM AvD5, 60 mM KCl, 10 µM EDTA)

To further characterize stabilizing and destabilizing effects of the pH in the catalytic triad region and the bulge, respectively, indicated in the NMR spectra, we performed UV melting studies under the same ionic conditions as the NMR experiments. Thermal denaturation between 20 and 90°C was followed through the absorbance at 260 nm at three different pH values (Figure 2.12). The curves show only a single melting transition at $T_m \sim 63°C$ which has to be an overlay of at least two helical regions, making the evaluation of small changes difficult. Furthermore, the curves are an overlay of the suspected stabilization at the catalytic triad and destabilization at the bulge, two effects that can be expected to counteract each other as they occur on two ends of the same helix (helix 1). This might explain why we observe at most a very slight stabilization at pH 5.5 compared to pH 6.5 and 7.5: the melting temperature increases by only $\sim$1°C (Figure 2.12).

Taken together, even though UV melting studies are unable to confirm the evidence, NMR spectra indicate a stabilization in the catalytic triad region at low pH, while in the region of the bulge broadening, a possible exchange cross peak to G11H1 and the absence of observable imino protons for U10, G26 or U27 indicate a less stable region.

2.2.5 Protonation of A32N1 goes along with a conformational switch

The fact that catalytic triad peaks can be observed both at high and at low pH but not in between (compare Figures 2.2 and 2.7) indicates the conversion from one relatively stable conformation to another. What clues do we have about the nature of these conformations? Unfortunately, since many of the NOE cross peaks are still to some degree broad also at pH 5 and pH 9, precise quantification of peak intensities is difficult and smaller changes in inter-proton distances can not be picked up. However, at least at low pH several characteristic NOE cross peaks appear in the spectra such as G4H8-A32H2 and a comparatively strong G4H8-H1' (Figure 2.7 top). Comparing this evidence with the variety of possible G-A base pairs[108,337,339] depicted in Figure 2.13, our data strongly suggests a G4$_{syn}$-A32H$^{+}_{anti}$ pair as in Figure 2.13c at low pH, where A32N1 is protonated. The *syn*-conformation causes an intense G4H8-H1' peak and A32H2 would be very close to G4H8 only in this form of base pair.

The imino resonance at 11.1 ppm assigned to G4H1 is similar to literature values for DNA G$_{syn}$-AH$^{+}{}_{anti}$ base pairs[340,341]. In a G$_{syn}$-AH$^{+}_{anti}$ base pair the guanine amino group can form a stabilizing hydrogen bond to the 5' phosphate[342]. This could decrease the rate of exchange of the two amino protons, causing the disappearance of the strong averaged amino peak observed at neutral pH for G4 (see Section 2.1.4). The adenine amino group is involved in a hydrogen bond to O6 of the guanine which leads to strong and well separated amino proton cross peaks (see A32H61/H62 in red in Figure 2.10a). Their positions are also very comparable to the ones observed in DNA G$_{syn}$-AH$^{+}_{anti}$ base pairs[340,341,343].

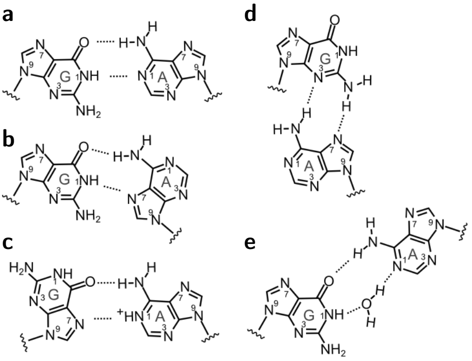


Figure 2.13 Variability of G-A base pairs[337]. a) G$_{anti}$-A$_{anti}$, b) G$_{anti}$-A$_{syn}$, c) G$_{syn}$-AH$^{+}_{anti}$, d) sheared G-A (N7-amino, amino-N3)[338], e) G$_{anti}$-A$_{anti}$ with water bridge[108].

The conformation of the G-A base pair at physiological to high pH is less amenable than the low pH conformation due to the lack of information on imino protons and the lack of characteristic NOE cross peaks that would point to a specific conformation. G4H8-H1' is less intense than at low pH and also the other available information is consistent with a G4 *anti*-conformation. However, it is not possible to determine unequivocally if G4 is stacking within the helix at high pH without observable protons on the Watson-Crick side of the base and with only rather weak distance restraints inferred from the broadened NOE cross peaks. In contrast, the strong inter-strand cross peak of the non-exchangeable H2 on the W-C face of A32 to C5H1' is good evidence for stacking of A32 inside the helix both at high and at low pH.

2.3 Solution structure

The structure of AvD5 proved more difficult to access than expected due to the dynamics at different time scales in loop, bulge and catalytic triad. We calculated two sets of structures — one for pH 5.2 and one for pH 7.8 — that differ only in the restraints used in the region of the catalytic triad. We used only the high-pH restraints in the bulge region, because we can not observe its peaks at low pH, and in all remaining regions, because they do not change significantly with pH.

Experimental restraints in bulge and catalytic triad are not sufficient to restrict more long-range bending and twisting motions between the adjacent helical parts. The introduction of 26 residual dipolar coupling (RDC) restraints narrows the accessible conformational space considerably (see Figure 2.14 and Table 2.3), but still the ensemble allows for some bending and twisting and is more well-defined only locally.

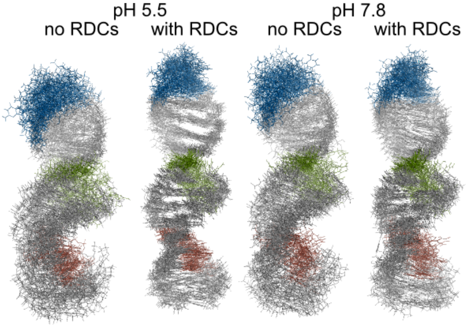


Figure 2.14 AvD5 structure ensembles calculated from NMR data collected at pH 5.5 and pH 7.8 and in the absence or presence of RDCs. The 20 lowest energy structures out of 200 calculated ones are shown.

The structures appear much more elongated than the conformations found in the crystal structures[106,236,237,280,284]. The same is observed also in the solution structure of D5 from Sc.ai5γ (ScD5)[281] and — even if much less pronounced — in D5 from Pl.lsu/2 (PlD5)[282]. Partly such global differences might be ascribed

Table 2.3 Structure calculation statistics for AvD5. The two sets of restraints at pH 5.2 and 7.8 differ only in the catalytic triad region. In the bulge, peaks are not observable at low pH, therefore restraints collected at high pH were used. In all other regions NOEs do not change significantly with pH.

	pH 5.2		**pH 7.8**	
	without RDCs	with RDCs	without RDCs	with RDCs
RMSDs[a] *(Å)*				
overall (1-35)	5.29 ± 1.84	2.65 ± 0.62	4.85 ± 1.55	3.43 ± 0.91
catalytic triad (3-5,31-33)	0.91 ± 0.29	0.60 ± 0.22	2.17 ± 0.89	2.35 ± 0.82
helix 1 (5-8,28-31)	0.52 ± 0.16	0.77 ± 0.25	0.51 ± 0.13	0.64 ± 0.18
bulge (9-10,24-27)	2.88 ± 0.77	2.45 ± 0.79	2.61 ± 0.96	2.00 ± 0.81
helix 2 (11-14,20-23)	0.45 ± 0.14	0.49 ± 0.20	0.44 ± 0.12	0.53 ± 0.21
loop (15-19)	1.25 ± 0.35	1.50 ± 0.36	1.20 ± 0.33	1.64 ± 0.66
distance restraints	700	700	660	660
intranucleotide (i=j)	244	244	238	238
sequential (|i-j|=1)	328	328	310	310
medium (2≤|i-j|≤5)	14	14	14	14
long range (5<|i-j|)	114	114	98	98
dihedral restraints	256	256	256	256
H-bond restraints	63	63	56	56
RDC restraints	-	26	-	26

[a] Root mean square deviations (RMSDs) for the ensembles of 20 lowest energy structures calculated in Xplor-NIH are given.

to real changes experienced by D5 when docked into the intron core or stabilized by crystal lattice contacts. There is no doubt that there are large changes in the local bulge structure, which can be expected to affect also the orientation of the adjacent helices. However, recent studies indicate that very elongated structures could be a calculation artefact, which can be remedied by the use of a more sophisticated force field, including explicit solvent and a proper treatment of electrostatics[279].

2.3.1 Loop

The U-rich pentaloop is the region of AvD5 with the highest proton density and therefore the highest density of distance restraints, mainly due to the nucleobase of A16 which is nicely embedded at its center (Figure 2.15). Because there is strong evidence for fast conformational exchange involving the G17 nucleobase (see Section 2.1.3), no NOE restraints involving G17H8 were incorporated in the calculations, making it relatively free to move in a solvent-exposed position. Also U18 is very solvent-exposed, while U19 is stacked into the loop and alternating between an interaction to U15 and A16, both of which are stacked on top of helix 2. Incorporation of G17 and U18 RDCs does not improve convergence of the local conformation of the loop, which is very likely due to the conformational averaging, and RDCs of G17 and U18 were therefore excluded from the calculations.

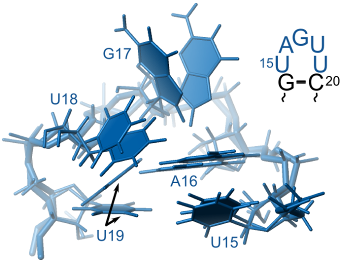


Figure 2.15 Conformation of the UAGUU pentaloop of AvD5 in the two lowest energy structures of the structure ensemble. Loop residues shown on the right are marked in blue in the secondary structure scheme. Black arrows mark the two distinct conformations of U19, bonding with either U15 and A16, respectively.

2.3.2 Bulge

Bulge nucleotides U10 and A24 are found in a stacked-in position in all structures of the AvD5 NMR ensembles (Figure 2.16). With very few exceptions where it extrudes into the minor groove, also C25 stacks in under A24. U10 is seen to interact with A24 or C25 or just situated loosely below or above C25. In many cases C25 seems to interact with the amino group of A9 (Figure 2.16a). G26 is bulged out either into the minor or into the major groove (compare Figure 2.16a and b). Both conformations are equally populated in the ensemble of calculated structures. We can not exclude either of them on the basis of violations of any experimental restraints. Also restraining the glycosidic angle of G26 to the *syn*-range — which we did not do for the final structure ensembles (see Section 2.1.2) — does not discriminate against one of the two conformations. Nonetheless, neither the co-existence of the two species — we observe only one set of peaks — nor a fast exchange between the two conformations — the passage from the major into the minor groove requires a considerable rearrangement — seems very likely. Very probably one of the two conformations would be suppressed if it was possible to collect more experimental restraints for this region.

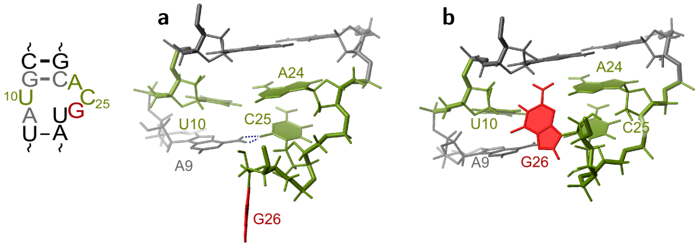


Figure 2.16 The structure of the AvD5 bulge. Residues are coloured according to the secondary structure scheme on the right, bulge nucleotides U10, A24 and C25 green, G26 red and the adjacent G11-C23 base pair as well as A9 are grey. a) A representative structure of AvD5 with G26 bulged into the major groove. Interactions of C25 with A9 are indicated. b) A representative structure of AvD5 with G26 bulged into the minor groove.

2.3.3 Catalytic triad

The structure of the catalytic triad at pH 5.2 is very well defined due to the stabilization of the A3-U33 base pair and formation of a G4-A32H$^+$ pair. In contrast, the scarce experimental restraints for G4 at pH 7.8 do not keep it from exploring a wide conformational space (Figure 2.17).

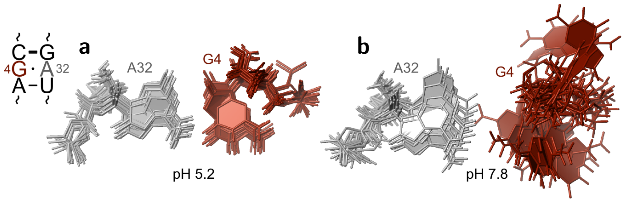


Figure 2.17 Conformation of the G-A base pair in the catalytic triad of AvD5 at acidic (a) and basic (b) pH. The structures were overlaid on all atoms of the catalytic triad except G4 (residues A3, C5, G31, A32 and U33; G4 was excluded because of the strong disorder at high pH.

2.4 Bulk fluorescence studies indicate pH dependent unstacking and Mg(II) binding in the bulge region

Protonation at the bulge seems to occur at A9. Nonetheless, protonation affects the nucleotides on the opposite strand of the bulge (A24, C25, G26) particularly strongly and resonances disappear quickly when the pH is lowered beyond the pK_a. While the structural rearrangement in the catalytic triad can be followed through the appearance of new NOE cross peaks, NMR spectroscopy does not help us in the case of the bulge region. We therefore decided to look at it from another angle, by employing the fluorescent nucleotide analogue 2-aminopurine (AP, Figure 2.18). AP nucleotides can be incorporated into nucleic acids and form base pairs with uracil that are isosteric to A-U pairs[344]. AP fluorescence has a maximum at 372 nm and is quenched to a large extent when it is stacked inside a helix. Structural rearrangements that change the degree of stacking can therefore be followed through fluorescence spectroscopy.

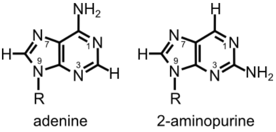


Figure 2.18 Adenine and its fluorescent analogue 2-aminopurine.

In AvD5 we see A24 firmly stacked onto helix 2 at pHs above the pK_a through cross peaks from A24H2 to the opposite strand. These cross peaks get lost below the pK_a along with the other bulge peaks and so can not tell us about the nature of structural changes in the bulge. The disappearance is not necessarily a sign of a movement of the A24 base, it could be due only to the broadening effect of conformational exchange not directly involving A24. By replacing A24 with an AP nucleotide (AvD5-AP24 construct) we wanted to find out if this nucleotide becomes more solvent exposed when the pH is lowered towards the pK_a. Indeed the fluorescence intensity at 372 nm does go up with decreasing pH (Figure 2.19). The effect is relatively small — roughly 20-30% — between pH 7 and 6.5 and if there is an inflection point close to the pK_a value of 6.4 which we determined in the bulge region by NMR spectroscopy, it is masked by the more dramatic fluorescence increase below pH 6.5. Thermal denaturation of RNA containing AP nucleotides, for instance, leads to a 3- to 10-fold increase of fluorescence intensity[345] and a similar increase might be expected upon acid denaturation. In the absence of a control it is thus difficult to judge how much of our final fluorescence increase at pH 5 could be specific to a conformational change in the AvD5 bulge — A24 or C25 could have pK_as around 5-5.5, which we can not measure because everything in the bulge becomes unobservable in the NMR spectra — and how much is just a result of the beginning denaturation of the adjacent helices in acidic conditions. However, the increase between pH 7 and pH 6.5 is so small that a rearrangement of A24 to a completely solvent exposed position that is linked to the protonation at A9-N1 and its pK_a of 6.4 can be excluded.

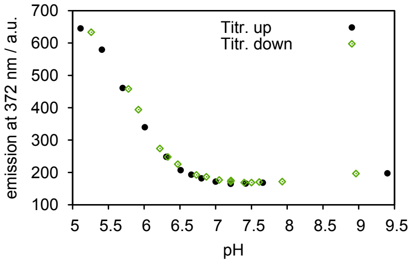


Figure 2.19 Fluorescence emission at 372 nm in AvD5-AP24 upon pH titration. Two titrations are plotted, one starting at low pH going up and a second one starting at high pH and going down. (1 µM AvD5-AP24, 60 mM KCl, 10 µM EDTA)

While the pH-dependence of the bulge does not seem to involve a significant rearrangement of the A24 base, it could instead induce changes in the sugar-phosphate backbone of the region, which is known to be important for the coordination of two Mg(II) ions in the assembled group II intron[106,236]. To find out if the pH influences Mg(II) binding to the bulge we titrated AvD5-AP24 with $MgCl_2$ at pHs above and below the previously determined pK_a in the AvD5 bulge (see Section 2.2.1).

We now performed our experiments at varying pH at 10 mM KCl to compare with the previous results and at 60 mM KCl to reproduce the conditions of our NMR experiments (Figure 2.20). While monovalent ions alone — as shown in the previous study[346] — do not lead to increased AP fluorescence, we find that they have a strong competition effect on Mg(II) binding. The more KCl is present, the weaker is the initial increase in intensity and also the lower is the apparent Mg(II) affinity, with K_Ds changing by more than an order of magnitude (Figure 2.20 and Table 2.4). K_D differences are less pronounced but the same general trend is observed with increasing pH, i.e. with decreasing H^+ concentration.

The curves in Figure 2.20 suggest that Mg(II) initially leads to a more solvent-exposed position of AP24 in low ionic strength and at low pH than in high ionic strength or at high pH. However, after the initial increase the fluorescence decreases again in most of the curves. The effect is more pronounced at lower pH and lower monovalent ion concentration. All the titrations end with relatively high ionic strength (100 mM $MgCl_2$), which could be the reason why all the curves seem to converge towards the end. At such high Mg(II) concentrations the influence of H^+ and K(I) concentrations on the conformation of the bulge

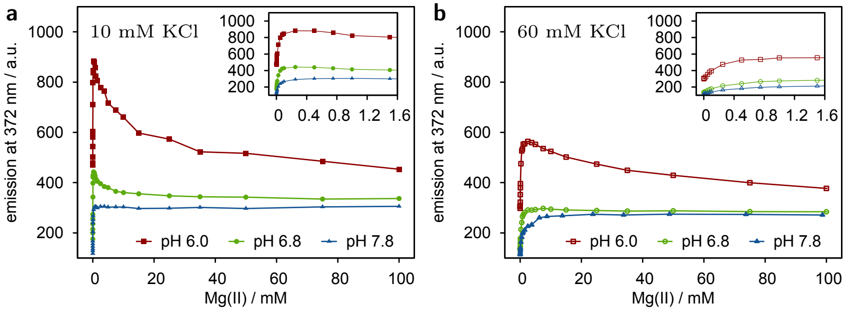


Figure 2.20 Fluorescence emission at 372 nm in a 1 µM AvD5-AP24 sample upon $MgCl_2$ titration in 10 mM MOPS buffer, 12.5 mM DTT and 10 mM (a) or 60 mM (b) KCl. The small insets show a zoom of the first few data points of the same curves.

Table 2.4 Dissociation constants determined for the bulge region of AvD5 by fitting AP24 fluorescence increase. Details of the fits are given in Appendix 13.

KCl [mM]	MOPS [mM]	DTT [mM]	$K_{D,pH6.0}$ [µM]	$K_{D,pH6.8}$ [µM]	$K_{D,pH7.8}$ [µM]
10	10	12.5	18.11 ± 2.16	20.33 ± 1.49	32.75 ± 2.60
60	10	12.5	200.46 ± 13.71	265.19 ± 13.66	817.26 ± 89.03

might be diminishing.

The origin of the strong decrease at higher Mg(II) concentrations remains puzzling, but it could lie in real conformational changes experienced by the bulge. While one strongly binding Mg(II) favours a clearly more unstacked AP24, higher divalent ion concentration could cause a compaction of the bulge structure in later stages that is responsible for the subsequent decrease in fluorescence.

In a previous study on ScD5 the strong decrease was attributed to oxidation[346], promoted by the mixing with the pipet at each $MgCl_2$ addition and could be alleviated, but not abolished, by the addition of DTT as an oxygen scavenger. In our experiments we observe a strong decrease even in the presence of 12.5 mM DTT. The still significant decrease might have been less apparent in the previous study simply because it did not test such high Mg(II) concentrations (the maximum concentration was 7 mM $MgCl_2$). The affinity constant determined in the ScD5 bulge was 7.9 ± 0.2 µM at first and after HPLC-purification it had changed to 196.0 ± 9.9 µM. In light of our results now this could indicate that the salt concentration in the samples was very different before and after HPLC.

2.5 Interactions with di- and trivalent metal ions characterized using NMR spectroscopy

Mg(II) is the natural co-factor of RNA and thus the most interesting metal ion in relation with RNA and with D5 in particular, since there is evidence for Mg(II) binding both in solution[75,256,266,269,281,282] and in the crystal structure[106,236,237]. We thus looked at the chemical shifts and line broadening of ^{1}H, ^{15}N and ^{13}C resonances upon addition of Mg(II). However, chemical shifts are influenced not only by the electron density of a binding metal ion but also by structural changes that might occur upon a distant binding event or merely as a response to the increasing ionic strength. For less ambiguous signals, Mn(II) and cobalt(III)hexammine were employed as Mg(II) mimics. Mn(II) leads to strong paramagnetic line broadening in the regions where it is binding. Cobalt(III)hexammine's exchange-inert NH_3 ligands yield observable NOEs to RNA protons at binding sites.

2.5.1 Magnesium(II)

^{1}H chemical shift perturbations

We performed titrations in the range of 0.5 to 12 mM $MgCl_2$ and followed the chemical shift changes ($\Delta\delta$) in ^{1}H,^{1}H-NOESY spectra. The absolute changes were evaluated at a 1:5 ratio of RNA to Mg(II) (Figure 2.21) because at higher concentrations some peaks are not observable anymore. Even if the shifts are much more pronounced with Mg(II), the overall picture is remarkably similar to the effects of KCl, i.e. peaks that are most strongly influenced by Mg(II) also shift most strongly and in the same direction with KCl (compare Figure 2.21 with Appendix 2b). In both cases the strongest chemical shift changes are observed in the bulge region.

Many titration curves at the 5' end and in and around the bulge region are distinctly biphasic (Figure 2.22a-c), indicating an overlay of either two binding events or structural changes. In the loop peaks barely move at all during the first

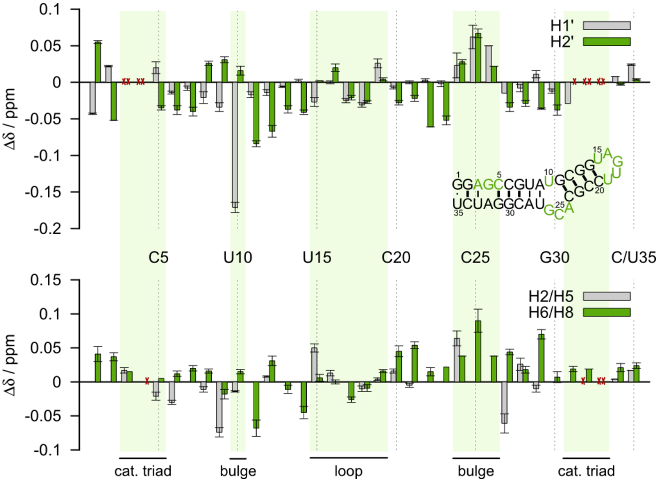


Figure 2.21 ^{1}H chemical shift changes upon addition of 5 equivalents of $MgCl_2$. Averages of four individual $MgCl_2$ titrations are plotted. To distinguish resonances that are not moving at all from values that are missing because assignment is not possible under all conditions the latter were crossed out red in the graph. (0.6-1.4 mM AvD5, 290 or 300 K, pD 6.7, 60 mM KCl, 10 µM EDTA)

6-7 equivalents of $MgCl_2$, then a quite linear increase starts where no plateau is reached at the highest tested Mg(II) concentration of 12 mM (corresponding to maximally 20 RNA equivalents, Figure 2.22d). Therefore, while the loop resonances are the least perturbed in D5 in the initial steps of the titration (Figure 2.21), they shift more strongly at higher Mg(II) concentration. The initial lag phase can at least partly be accounted for by the initial sequestration of the bulk Mg(II) to more favourable binding sites in D5 (see the analysis of the binding affinities below), while the later increase without plateau could be a sign of weak, non-specific interactions as the free Mg(II) concentration goes up.

Interestingly, in Figure 2.21 the H2' in the helical regions are always affected more strongly than the H1' and always move upfield, with the exception only of U8 and A9 next to the bulge. This could be attributed to favourable but non-specific electrostatic interactions of Mg(II) at the O2' sites. However, K(I) does not have such a particular effect on the H2' resonances (Appendix 2).

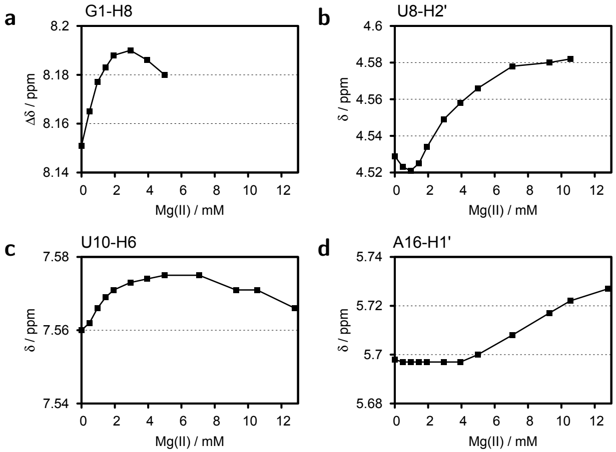


Figure 2.22 a) Chemical shift changes at the 5' end and b,c) in the AvD5 bulge show a biphasic behaviour when plotted against increasing $MgCl_2$ concentration. d) Chemical shifts in the loop start to change significantly only after several equivalents of Mg(II) have been added.

Chemical shift perturbations of the heteronuclei

Both the aromatic ^{13}C and the ^{15}N shifts (Appendix 14) indicate strongest Mg(II) interaction at the 5' di- or triphosphate. In the rest of the molecule, the aromatic ^{13}C chemical shift perturbations basically follow the ^{1}H ones with strongest changes in the bulge and negligible changes in the loop. Regarding the nitrogens, apart from 5' end resonances, only the N7 resonances of G13/G14 and G30/G31 seem to be influenced moderately by the addition of Mg(II), which can probably be attributed to a favourable interaction at the major groove edges of the two Gs of the tandem G-C pairs[347].

Apparent Mg(II) affinity constants

Metal ion affinities can be derived by fitting chemical shift perturbations as a function of Mg(II) concentration to a 1:1 binding isotherm (Equation 4.3 in the Experimental Section). This requires that curves which are subject to more than one influence (more than one metal ion or additional structural changes) can be excluded confidently. However, even if these curves can be recognized and excluded, K_D values obtained from the remaining curves are only apparent affinities when there is more than one binding site in the molecule because the actual free metal ion concentration is reduced by each binding site. If it is possible to determine the number of binding sites in a molecule and localize them, initial K_D estimates can be used in an iterative procedure to determine the number of metal ions already bound to other sites[114,145]. In this way in principle the real intrinsic affinities can be established (see next section).

We plotted the chemical shift perturbations against Mg(II) concentration for all ^{1}H resonances assigned in NOESY spectra. Curves that move less than 0.02 ppm, curves that have too few data points, e.g. due to broadening of the peaks, and curves that are clearly biphasic were excluded, all others were fit to Equation 4.3 for a single binding isotherm to determine initial, apparent K_Ds ($K_{D,ini}$ in Table 2.5 and black curves in the graphs in Appendix 15). Those were used as the basis to group individual protons into binding sites in AvD5 (see next section).

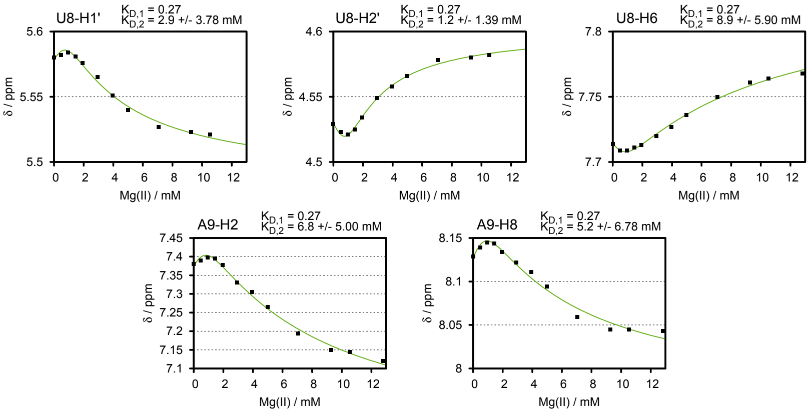


Figure 2.23 Mg(II)-dependent chemical shift changes in the bulge of AvD5 can be fit to an equation for two independent binding sites (Equation 4.4). Guided by the results of the fluorescence studies (Section 2.4) the first K_D was restricted to 0.27 mM. Without this restriction, no reasonable fits were obtained.

An equation for two independent binding events (Equation 4.4) was used to fit the curves displaying biphasic behaviour, mainly to check the consistency of the NMR and fluorescence data (Section 2.4). The number of titration data points in the NMR experiments is insufficient for an accurate five-parameter fit which would be required to determine both K_Ds. However, if the first K_D is restricted to the range around 0.27 mM as determined by fluorescence measurements (at

NMR conditions: 60 mM KCl, pH 6.8 in Table 2.4) several of the biphasic curves in the bulge region can be fit reasonably well, indicating a second apparent K_D in the low millimolar range, which is the range also obtained from monophasic curves of nearby resonances (compare $K_{D,2}$ in Figure 2.23 to $K_{D,ini}$ in the bulge region in Table 2.5).

All determined apparent K_Ds are in the low to intermediate millimolar range. K_Ds around the bulge are comparable to the ones found in the bulge of ScD5 in a previous study[281]. In the loop the curves can not be fit because of the initial lag before the chemical shifts change noticeably (Figure 2.22d).

Intrinsic Mg(II) affinity constants

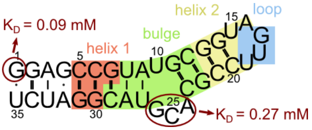


Figure 2.24 Mg(II) binding sites assigned on the basis of similar apparent K_D values. K_Ds for two binding sites at the 5' end and in the bulge (circled in red) are known from the literature[114] and from our fluorescence experiments (Section 2.4), respectively. Four additional binding sites are indicated by different colors.

Erat et al.[114,145] developed an iterative fitting procedure (implemented in the Matlab script ISTAR) that allows to determine the intrinsic affinities of multiple metal ion binding sites within one molecule. It requires that all binding sites without exception can be identified, because this information is used to iteratively calculate the actual free metal ion concentration, which is the basis for the determination of the intrinsic affinities. In AvD5, we already know about two stronger binding sites: one at the 5' end (K_D = 0.09 mM inferred from the published Mg(II) affinity for the triphosphate at the 5' end of an RNA[114]) and one at the bulge, determined in our fluorescence study (K_D = 0.27 mM, see Table 2.4) (Figure 2.24). These two K_D values were incorporated into the calculation as they are, even if the K_D for the bulge is only an apparent value and the intrinsic one would be lower.

However, unfortunately the unambiguous assigment of the rest of the Mg(II) binding sites in AvD5 proved far from trivial. ISTAR[145] depends on the grouping of protons into a limited number of binding sites based on closeness in space and similar K_Ds. In AvD5 it is possible to define four such binding sites (helix1, bulge, helix2 and loop in Figure 2.24 and Table 2.5). These four binding sites are in addition to the stronger ones at the bulge and the 5' end, see previous

Table 2.5 Mg(II) dissociation constants as determined from proton chemical shift changes in ^{1}H,^{1}H-NOESY spectra using Equation 4.3 (see Experimental Section). $K_{D,ini}$: apparent K_D values obtained by plotting against the total added Mg(II). $K_{D,fin,6sites}$ and $K_{D,fin,7sites}$: estimates of the intrinsic K_D values obtained by the iterative fitting procedure of ISTAR [145] assuming a total of six and seven binding sites, respectively (Figure 2.24, binding sites are indicated in the same colors as in this table). Many K_D values are missing for the loop region because an initial lag phase of the curves prevents reasonable fits. The values in parentheses give the error in %. K_Ds of resonances which were excluded from the iterative fitting procedure due to various reasons (see text) are listed in Table 2.6. All fitting curves are shown in Appendix 15.

	$K_{D,ini}$	$K_{D,fin,6sites}$	$K_{D,fin,7sites}$
bulge			
G11-H2’	6.66 ± 2.23 (33.46%)	0.99 ± 0.21 (20.81%)	0.65 ± 0.13 (19.57%)
G11-H8	7.42 ± 1.70 (22.96%)	1.72 ± 0.29 (16.59%)	1.15 ± 0.16 (13.53%)
C12-H6	5.25 ± 1.37 (26.05%)	1.41 ± 0.10 (6.89%)	0.96 ± 0.09 (8.85%)
G22-H8	4.31 ± 1.16 (26.86%)	1.19 ± 0.18 (14.85%)	0.81 ± 0.12 (14.39%)
C29-H2’	3.25 ± 0.57 (17.51%)	0.98 ± 0.25 (25.78%)	0.67 ± 0.17 (24.77%)
C29-H6	4.42 ± 0.80 (18.05%)	1.22 ± 0.13 (10.37%)	0.84 ± 0.09 (10.34%)
helix1			
C5-H2’	9.56 ± 1.25 (13.03%)	2.71 ± 0.59 (21.62%)	1.79 ± 0.40 (22.05%)
C6-H1’	16.85 ± 7.03 (41.69%)	4.08 ± 0.51 (12.55%)	2.66 ± 0.28 (10.62%)
C6-H5	10.88 ± 4.79 (43.99%)	2.89 ± 0.63 (21.72%)	1.89 ± 0.36 (19.15%)
G7-H2’	15.07 ± 2.50 (16.56%)	4.08 ± 0.65 (16.06%)	2.65 ± 0.41 (15.30%)
G7-H8	12.62 ± 4.12 (32.65%)	3.11 ± 0.33 (10.58%)	2.04 ± 0.22 (10.98%)
G30-H2’	16.23 ± 4.70 (28.97%)	3.90 ± 0.85 (21.79%)	2.57 ± 0.60 (23.42%)
G31-H8	8.35 ± 1.94 (23.23%)	2.26 ± 0.45 (19.86%)	1.48 ± 0.29 (19.82%)
helix2			
G13-H2’	12.23 ± 3.50 (28.63%)	3.07 ± 0.49 (15.85%)	1.99 ± 0.27 (13.70%)
G14-H2’	25.34 ± 7.64 (30.17%)	5.42 ± 0.48 (8.87%)	3.50 ± 0.28 (8.08%)
G14-H8	16.81 ± 4.75 (28.23%)	3.87 ± 0.29 (7.44%)	2.51 ± 0.17 (6.83%)
U15-H1’	43.65 ± 13.66 (31.29%)	7.56 ± 1.10 (14.57%)	4.81 ± 0.75 (15.64%)
A16-H2	20.35 ± 6.78 (33.32%)	4.34 ± 0.49 (11.17%)	2.82 ± 0.36 (12.61%)
A16-H3’	14.29 ± 2.86 (20.00%)	3.61 ± 0.91 (25.18%)	2.37 ± 0.61 (25.65%)
U19-H6	11.38 ± 4.15 (36.48%)	2.62 ± 0.36 (13.65%)	1.70 ± 0.21 (12.64%)
C21-H2’	10.74 ± 2.37 (22.07%)	2.77 ± 0.51 (18.27%)	1.80 ± 0.31 (17.41%)
C21-H6	50.05 ± 26.08 (52.11%)	8.02 ± 0.53 (6.67%)	5.07 ± 0.28 (5.57%)
loop			
G17-H1’	-	16.47 ± 2.90 (17.62%)	9.09 ± 0.90 (9.86%)
G17-H2’	-	27.87 ± 10.64 (38.18%)	13.50 ± 3.05 (22.61%)
G17-H8	-	59.22 ± 34.48 (58.22%)	22.09 ± 5.28 (23.89%)
U18-H1’	-	-	92.73 ± 70.47 (76.00%)
U18-H2’	-	-	52.00 ± 30.02 (57.74%)
U18-H3’	-	-	73.43 ± 64.57 (87.94%)
U18-H4’	-	25.70 ± 10.95 (42.61%)	11.03 ± 2.90 (26.29%)
U19-H1’	-	-	57.11 ± 52.61 (92.11%)
C20-H5	-	19.69 ± 4.84 (24.59%)	10.47 ± 1.51 (14.43%)

Table 2.6 Mg(II) dissociation constants as determined from ^{1}H chemical shift changes in AvD5, continued. The table shows only the values that were excluded from the iterative fitting procedure in ISTAR[145]. See legend of Table 2.5 for further information.

	$K_{D,ini}$	$K_{D,fin,6sites}$	$K_{D,fin,7sites}$
excluded			
A3-H2	5.41 ± 1.32 (24.30%)	0.48 ± 0.20 (40.78%)	0.25 ± 0.13 (53.29%)
C6-H2'	6.10 ± 1.87 (30.69%)	0.96 ± 0.15 (15.96%)	0.54 ± 0.09 (17.51%)
A9-H1'	12.78 ± 5.06 (39.56%)	2.50 ± 0.23 (9.25%)	1.50 ± 0.11 (7.53%)
U10-H1'	7.41 ± 1.58 (21.25%)	1.57 ± 0.26 (16.62%)	0.95 ± 0.18 (18.96%)
U10-H5	32.29 ± 26.56 (82.25%)	5.33 ± 1.34 (25.16%)	3.20 ± 0.66 (20.46%)
G11-H1'	14.33 ± 5.98 (41.72%)	2.97 ± 0.54 (18.09%)	1.82 ± 0.31 (16.82%)
C12-H2'	3.14 ± 1.08 (34.48%)	0.52 ± 0.13 (24.63%)	0.28 ± 0.07 (26.02%)
G13-H8	-	-	49.71 ± 51.10 (102.80%)
G14-H3'	-	-	29.78 ± 9.89 (33.22%)
U15-H5	-	11.56 ± 1.48 (12.83%)	6.58 ± 0.81 (12.27%)
U15-H6	-	-	-
A16-H2'	-	-	28.85 ± 12.81 (44.39%)
U18-H5	-	-	-
C20-H6	-	10.99 ± 1.16 (10.59%)	6.24 ± 0.32 (5.17%)
A24-H2	5.76 ± 2.07 (35.94%)	0.41 ± 0.16 (40.03%)	0.19 ± 0.11 (58.09%)
U27-H2'	12.75 ± 2.81 (22.06%)	2.83 ± 0.46 (16.32%)	1.73 ± 0.28 (16.33%)
A28-H2	1.10 ± 0.35 (32.16%)	0.21 ± 0.23 (109.09%)	0.16 ± 0.21 (127.34%)
A28-H2'	10.38 ± 3.16 (30.44%)	2.21 ± 0.31 (14.09%)	1.33 ± 0.18 (13.17%)
G30-H1'	29.50 ± 22.46 (76.13%)	4.66 ± 0.99 (21.24%)	2.78 ± 0.51 (18.31%)
C35-H1'	38.10 ± 18.16 (47.67%)	6.52 ± 1.29 (19.84%)	3.91 ± 0.66 (16.79%)

paragraph. However, in this process a significant number of protons, whose K_Ds deviate a lot from that of their neighbours, has to be excluded (see Table 2.6). This rather large number of outliers indicates that the assumption that each proton is only influenced by one metal ion binding event, does not hold for many of the protons. This should be the case especially at the border between two binding sites, but also in the bulge, where resonances are expected to experience an additional influence from a conformational change (see Section 2.4). We can not confidently exclude the possibility that also the protons included in the calculation are subject to multiple influences. As a consequence, we have to assume a high uncertainty in the determined intrinsic K_Ds ($K_{D,fin,6sites}$ in Table 2.5).

Nonetheless, the results of the iterative fitting procedure allow a number of qualitative conclusions. Intrinsic K_Ds are about one order of magnitude smaller than the apparent ones. They are of comparable magnitude in the two helical regions and a little stronger in the bulge ($K_{D,fin,6sites}$ in Table 2.5). In the loop, K_D values are significantly higher and several curves can not be fit reasonably. The iterative procedure brings the first few data points closer together, because the estimated free metal ion concentration is decreased and thereby reduces the error of the least-squares fits for many curves, even if some can still not be fit (values that are not given in the loop region in Table 2.5). If we add a hypothetical seventh binding site in the calculation (we used $K_D = 1$ mM), the initial stagnation of the chemical shift can be accounted for much better ($K_{D,fin,7sites}$ in Table 2.5 and Figure 2.25). We thus have indirect evidence for one additional binding site, which we can, however, not localize based on our data. It could for example, be hidden in the region of the catalytic triad where there is little information because only few resonances can be followed well in the spectra of the Mg(II) titration. Alternatively, there might be an additional binding site close to the bulge.

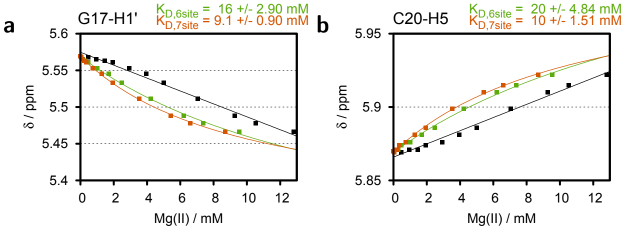


Figure 2.25 Mg(II)-dependent chemical shift changes in the loop region of AvD5. The initial lag phase in the chemical shift changes when they are plotted against the total added amount of Mg(II) (black curve) leads to a bad fit to a binding isotherm (Equation 4.3). When the Mg(II) that is sequestered by other binding sites is taken into account, much better fits can be obtained as seen in the green and orange curves, which are derived from calculations assuming a total of six and seven binding sites, respectively.

In summary, we propose 3-4 Mg(II) binding sites on the basis of NMR spectra with intrinsic K_Ds in the low millimolar range (in helix 1 close to the catalytic triad, in helix 2 close to the loop and in or close to the bulge) and another weaker binding site at the top of the pentaloop. These binding sites are in addition to another stronger binding site at the bulge determined in fluorescence experiments and the metal ion binding to the phosphates at the 5' end.

2.5.2 Cadmium(II)

Cd(II) is a softer metal ion than Mg(II) and thus has a stronger affinity for nitrogen ligands. It can be an advantage to employ Cd(II) as a substitute for Mg(II) to benefit from the tighter binding and observe ^{13}C and ^{15}N chemical shift changes in cases where Mg(II) binds more loosely and exchanges on the time scale of the chemical shift differences, leading to broadening of the signals. We titrated a ^{13}C,^{15}N-labelled AvD5 sample with 0.5 to 1.75 mM $Cd(ClO_4)_2$, following induced changes in ^{1}H,^{13}C-HSQC and ^2J ^{1}H,^{15}N-HSQC spectra. Appendix 14b shows the chemical shift changes at the first titration step of 0.5 mM Cd(II), because already at 1 mM Cd(II) many more resonances disappear completely. Like with Mg(II), the strongest shifts are observed at the two tandem G-C pairs (G13/G14, G30/G31) and at G1N7, only stronger and at lower metal ion concentration. G1N7 has disappeared already at 0.5 mM Cd(II), which we attribute to the strong binding and the formation of a macrochelate involving N7 and phosphates of G1[91,142]. We did, however, not titrate far enough to see the bound resonance of G1N7 appear like we did with D1kz (see Section 3.7.2).

2.5.3 Manganese(II)

The titration of AvD5 with $MnCl_2$ complements the data collected on Cd(II) and Mg(II). By looking at the distance-dependent line broadening at substoichiometric metal ion concentrations the effect of metal ion proximity can be studied without disturbance by structural rearrangement. We titrated AvD5 with 15 to 120 µM $MnCl_2$ and followed the induced changes in ^{1}H,^{1}H-NOESY spectra. By far the strongest broadening is apparent at the 5' end, where 15 µM are already enough to broaden resonances completely (Figure 2.26). Surprisingly U15H5 also broadens at a rather low 30 µM, while A16 follows only at 90 to 120 µM and the rest of the loop residues G17 to U19 do not broaden at all up to this Mn(II) concentration. It thus seems more likely that U15 is affected by a metal ion interacting at the upper part of helix 2 rather than at the loop, presumably at the two G-C basepairs that also interact with Mg(II) and Cd(II). Many residues in the helices and in the bulge start to disappear at 60 to 90 µM $MnCl_2$. Broadening in the bulge region occurs in the same Mn(II) range as in the two helices, indicating no particularly strong binding site for Mn(II) in this region. The only well-resolved resonance in the catalytic triad — A3H2 — is

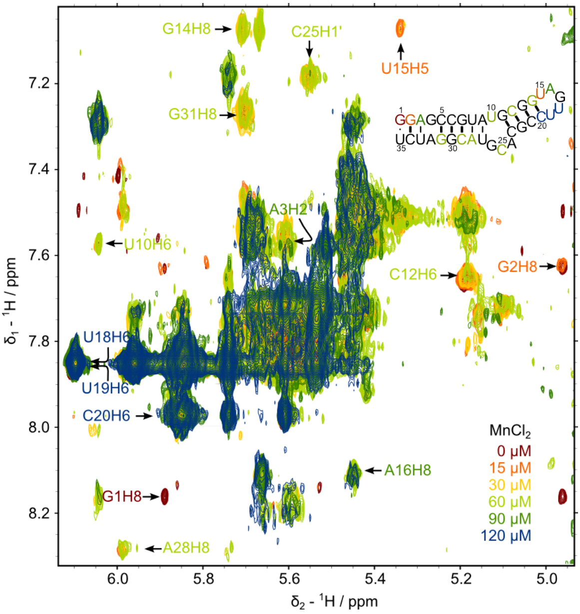


Figure 2.26 Overlay of AvD5 ^{1}H,^{1}H NOESY spectra with increasing Mn(II) concentration. Arrows mark well-resolved resonances that show broadening at various titration steps and the corresponding nucleotides are highlighted in the secondary structure inset in the colour of the last Mn(II) concentration, where they can still be observed. (0.67 mM AvD5, 290 K, pD 6.7, 60 mM KCl, 10 µM EDTA)

still well visible at 90 µM $MnCl_2$.

2.5.4 Cobalt(III)hexammine

While Cd(II) and Mn(II) serve mainly as probes for inner sphere coordination sites of Mg(II), the exchange-inert cobalt(III)hexammine complex can indicate probable binding sites for the hexaqua complex of Mg(II). We titrated cobalt(III)hexammine into an AvD5 sample up to a concentration of 2.4 mM. Initially, at 0.4 mM cobalt(III)hexammine (one RNA equivalent), only cross peaks to exchangeable protons in helix 2 are observed (to G22H1, G13H1, G14H1, U15H3 (see Figure 2.27a) and to amino protons of C12, C21 and C20). At 2x the RNA concentration, peaks to helix 1 resonances (G7H1, G30H1) and weak cross peaks to G13H8 and G14H8 start to appear. No strong cross peak is observed to G11H1 and U8H3, the observable resonances closest to the bulge. Increasing the cobalt(III)hexammine concentration further, almost all RNA resonances start to develop cross peaks to the bulk resonance of the complex, indicating unspecific binding. The most attractive sites for cobalt(III)hexammine seem to be the regions of the tandem G-C base pairs (G13/G14, G30/G31) in helix 1 and 2.

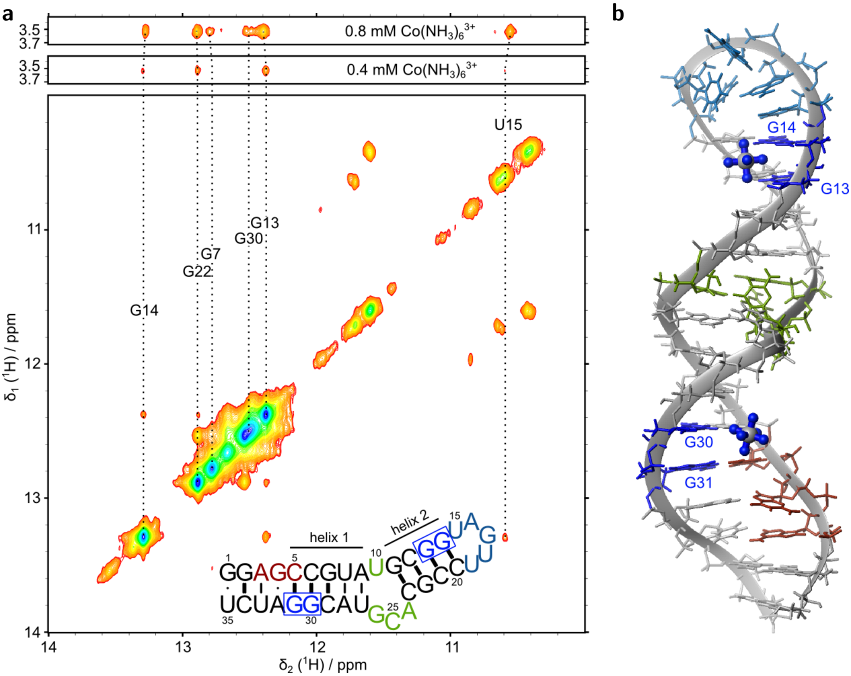


Figure 2.27 a) Imino proton region of a ^{1}H,^{1}H-NOESY spectrum in the presence of 0.4 mM cobalt(III)hexammine (*bottom*) and sections displaying the bulk resonance of cobalt(III)-hexammine ammonia groups at 3.6 ppm in the presence of 0.4 and 0.8 mM of the metal complex (*top*) (0.4 mM AvD5, 280 K, pH 6.7, 60 mM KCl, 10 µM EDTA). The two tandem G-C pairs which are the most favoured binding sites are highlighted in dark blue in the secondary structure inset, which indicates also catalytic triad (red), bulge (green), and loop (blue). b) AvD5 structure with two cobalt(III)hexammine molecules which were positioned by loose NOE-derived distance restraints. Residues are coloured according to the scheme in a).

The NOE information was incorporated as loose distance restraints in structure calculations using Xplor-NIH[348,349]. The restraints are satisfied well by two cobalt(III)hexammine molecules, which are positioned in the major groove close to the tandem G-C pairs (G13-G14 and G30-G31) (Figure 2.27b).

2.6 AvD5 - Discussion and conclusions

2.6.1 The G-A pair in the catalytic triad and its pH-dependent conformational switch

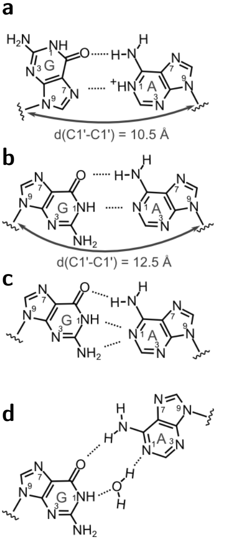


Figure 2.28 G-A base pairs with the adenine in *anti*-conformation. a) A G_{syn}-AH^{+}_{anti} is observed in the catalytic triad of AvD5 at low pH. b-d) G_{anti}-A_{anti} base pairs observed in the literature[108,339–341,343]. C1'-C1' distances are indicated.

We find that the G4-A32 mispair at the center of the catalytic triad forms a protonated G_{syn}-AH^{+}_{anti} pair (Figure 2.28a) below pH<6.7, but switches to a different conformation at higher pH. Such a conversion of a G-A base pair to protonated G_{syn}-AH^{+}_{anti} conformation at low pH has been described in the literature for DNA sequences[340,341,343] with reported pK_as of 5.9 to 6 (compare to pK_a 6.7 in the G4-A32 mispair in AvD5). We are not aware of any such study on RNA in solution.

In general, which and if one of the possible base pairs (Figure 2.13) is formed by a G-A pair depends on the sequence context (the nearest neighbours in particular)[350], but also on external factors like crystallization conditions, so that for example the very same DNA sequence can yield different conformations in solution and in a crystal[340,351–353].

Speculation on the G-A conformation adopted at high pH. The described DNA solution structures switch to an *anti-anti* base pair at neutral to basic pH[340,341,343] (Figure 2.28b). The base pair has been suggested to occur with a bifurcated H-bond of A-N1 to both G-N1 and G-N2[341] (Figure 2.28c), which is also observed in an RNA crystal structure by Leonard et al.[339]. A third similar, H_2O-mediated G_{anti}-A_{anti} base pair (Figure 2.28d) is proposed by Correll et al.[108] in the crystal structure of a 5S rRNA domain. A switching between those several closely related pairing possibilities could be the cause of broadened peaks in the catalytic triad of AvD5 which we observe even at pHs sufficiently far above the perturbed pK_a, which suggest that there is still some conformational averaging. The broadening might, however, also simply be a sign of a destabilized structure with no stable hydrogen bonds at all between G and A.

The *anti*-to-*syn*-conversion goes along with a distinct stabilization in certain DNA and RNA sequences[354–356]. Brown et al.[354] report an increase of the melting temperature from 19 to 38.5°C in a DNA duplex (in 100 mM sodium phosphate buffer), a much higher increase than what we observe in our UV melting experiments. We find only a very slight increase in melting temperature upon formation of the G_{syn}-AH^+_{anti} pair (see Section 2.2.4), even if NMR spectra indicate that the region is more stable under acidic conditions (see Section 2.2.4). The proximity of the destabilized bulge region might be the reason why this is not reflected in the melting temperature.

Single G-A mismatches and isostericity. Single mismatches inside a normal helix, if they are not isosteric to the canonical Watson-Crick base pairs[22], lead to distortions that influence their nearest neighbours, so that those frequently also do not form canonical base pairs[350]. In general G-A mismatches are therefore more common in terminal or internal loops or flanked by other non-canonical base pairs[108,357,358].

Davis et al.[350] compiled single-mismatches in available RNA 3D structures to correlate their conformations with the sequence of the flanking base pairs. While G-A pairs are the most common type of single mismatches, the nearest neighbour combination in AvD5 (5'AGU3'/5'GAU3') is not found frequently in the available RNA 3D structures[350]. In fact, a query of all presently available structures in the PDB[322] using MC-Search[360–362] yields only one result for this sequence element in a theophylline-binding RNA hairpin (PDB-ID: 1EHT[363], 1O15[359]). In this structure the A-U pair of the fragment is directly adjacent to an internal loop, is part of the binding platform for the small molecule and does not form a canonical Watson-Crick pair (Figure 2.29). The G-A pair is in an *anti/anti*-conformation, but the two bases are not co-planar. The C1'-C1' distance of a G_{anti}-A_{anti} pair is 12.7 Å (Figure 2.28b), which is sig-

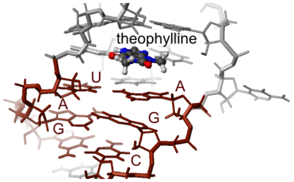


Figure 2.29 Another instance of the exact sequence of the catalyic triad region is found in the theophylline binding RNA (PDB-ID: 1O15[359]). In this structure the mispaired G and A do not bond and there is no W-C A-U pair.

nificantly larger than that of a canonical Watson-Crick pair (10.5 Å, see `http://rna.bgsu.edu/FR3D/basepairs/`[22]) and can disfavour the neighbouring A-U base pair, which might also be why we do not observe A3-U33 in AvD5 at high pH. In a G_{syn}-AH^{+}_{anti} base pair the distance is more similar to that in a Watson-Crick pair, explaining why A3-U33 can form stably at low pH (see Section 2.2.4).

The favourable decrease of the C1'-C1' distance to a value similar to the surrounding Watson-Crick base pairs might also be an important driving force in the low-pH conversion of the base pair. It might cause G4 to sample the *syn*-conformation even at higher pH, which could in turn induce the elevated $\mathrm{p}K_a$[15].

Single G_{syn}-AH^{+}_{anti} base pairs in published RNA structures. In general the reports of RNA solution structures with single G-A mismatches are few (PDB-IDs: 1P5M[336], 2HUA[364], 2KRL[365], 2EUY[366], results from MC-Search, excluding duplicates). To date (September 2012) there is only one solution structure containing a protonated G_{syn}-AH^{+}_{anti} (PDB-ID: 1P5M[336]) at pH 6.4, but no pH dependence was analysed in that study of domain II of the HCV IRES (Hepatitis C virus internal ribosome entry site). The sequence context in this case was 5'CGC3'/3'GAG5', suggesting that the C-G pair 3' of the G-A might predispose to this conformation, but this can not be confirmed in a MC-Search query for the combination of a G-A single mismatch next to a C-G base pair, where the majority of results has G and A not pairing at all or pairing via their Watson-Crick sides.

Using MC-Search and MC-Annotate[360–362,367,368] to query and analyse G-A Hoogsteen-W-C base pairs yields only a handful of structures, among which the above mentioned HCV IRES domain II[336] is the only one where the bases are positioned in a way to allow the two hydrogen bonds of the protonated G_{syn}-AH^{+}_{anti} to be formed. Two other structures have a G-A Hoogsteen-W-C base pair with only one hydrogen bond from A-N2 to G-N7 (PDB-IDs: 2Y17[369], 3OXB[370]). Maybe in these cases there is a protonated G_{syn}-AH^{+}_{anti} pair and it was overlooked because of resolution or because a possible protonation was not taken into account.

In summary, we describe here a rare single-mismatch G-A pair in RNA, for which the conformation could not have been predicted on the basis of any previous structural evidence. We observe for the first time in RNA a conformational switch from an *anti/anti*-conformation to a G_{syn}-AH^{+}_{anti} conformation. NMR spectra indicate that it goes along with some thermodynamic stabilization, even if the overlap of pH dependent effects in the bulge and catalytic triad regions does not permit to confirm a stabilization in UV melting studies.

2.6.2 Bulge structure and pH dependence

Comparative analysis of D5 bulge conformation

The D5 bulge changes its conformation not only with pH but also upon docking to the rest of the intron. To date the structure of this catalytically so crucial region has been solved in four different introns[106,236,238,280–282] (including this one). ScD5[281] (from Sc.ai5γ) and PlD5[282] (from intron Pl.lsu/2) have been studied in solution. There is also a crystal structure of ScD5 together with D6[280] and most recently the full group IIC intron from *O. iheyensis* has been solved[106,236–238,284], adding another D5 structure (OiD5). This structure is now available in several states thought to represent different stages of catalysis[106,238]. D5 remains principally the same in all of them apart from a minor reorientation of bulge nucleotide C25 (AvD5 numbering).

In all but PlD5, the sequence of the bulge is identical, including the A-U closing base pair at the bottom (A9-U27 in Figure 2.30a). PlD5 has a G-C closing pair below the bulge and an A instead of the bulged C25[282] (Figure 2.30a). Nonetheless, the structure of the bulge is much less different between PlD5 and the other D5s than between crystal and solution structures. In the crystals, a G-U pair is formed (corresponding to G26-U10 in AvD5), which stacks below the base pair at the top of the bulge (G11-C23 in AvD5) and the other two bulge nucleotides (A24 and C25 in AvD5) are looped out (Figure 2.30b). In contrast, the G corresponding to G26 in AvD5 is looped out in the two solution structures of ScD5 and PlD5 and A24 is always stacked below the previous base pair (Figure 2.30c). C/A25 is more free to move and sometimes extrudes into the minor groove, but usually it also stacks below A24. So in spite of sequence differences, both solution structures share main features which are now also found in our newly solved AvD5 (Figure 2.30d), even if base orientation and backbone slope vary slightly.

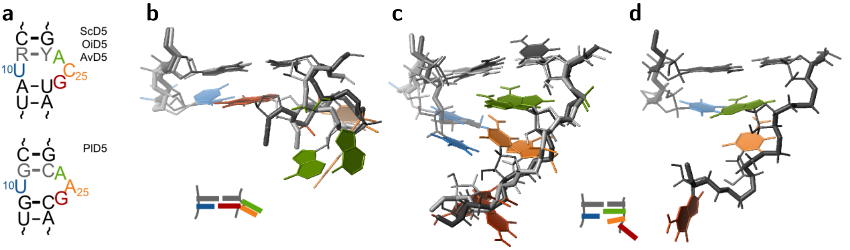


Figure 2.30 The D5 bulge in crystal and solution structures. a) The sequence of the bulge is identical in ScD5, OiD5 and AvD5, the only difference in PlD5 is A25. The colored bulge residues and the grey base pair on top are shown in a, b and c using the same coloring scheme. Residues are numbered to allow direct comparison to AvD5. b) The bulge structure in the crystal structures of D5 from Sc.ai5γ (PDB-ID: 1KXK[280]) and a group IIC intron from *O. iheyensis* (PDB-ID: 3G78[237]). b) The bulge in solution structures of ScD5 (PDB-ID: 1R2P[281]) and PlD5 (PDB-ID: 2F88[282]). c) A representative structure of the bulge in AvD5. The conformation of the bulge in the crystal and in solution is indicated schematically below the respective structures.

G26 is always looped out in the solution structures, but exclusively into the major groove in ScD5 and PlD5. Only in AvD5, G26 is bulged into the minor groove in about 50% of the calculated structures (see Section 2.3.2). We can not exclude those structures on the basis of violations of any experimental restraints, even if it is not likely that there is an actual structural difference between the bulge in ScD5 and AvD5. The structures where G26 is in the minor groove are even consistent with the published NOE restraints of ScD5[281], i.e. re-refinement of AvD5 with those restraints included does not bring up any violations. Nonetheless, neither the co-existence of the two species nor a fast exchange between the two conformations is probable considering the experimental results. We therefore conclude that one of the two conformations likely is an experimental artefact which would be suppressed if it was possible to collect more experimental restraints for the bulge region.

Structural basis of the pH dependence in the bulge

Elevated pK_as are favoured in structural contexts where the protonated base is stabilized. This can be achieved by a region of high phosphate density and therefore strong electrostatic potential. It can also be achieved by stabilizing base-base interactions, e.g. via hydrogen bonds[15]. The latter is strongly suggested to happen for A32 in AvD5, where the protonated position can form a new hydrogen bond to G4N7. However, in the case of A9 it is less obvious.

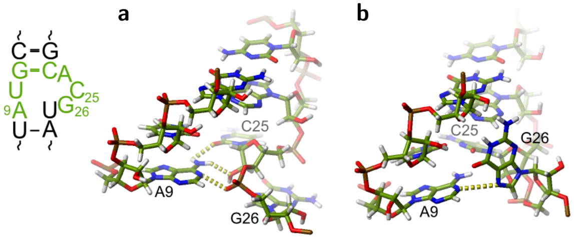


Figure 2.31 Possible polar and electrostatic interactions in the bulge region of AvD5 that could favour protonation of A9. The residues shown in a) and b) in two example structures from the AvD5 NMR ensemble are labelled green in the sequence scheme. The dashed yellow lines mark interactions mentioned in the text.

Considering the structure in the full group IIC intron crystal structure [236–238], where not only A9 is in a canonical Watson-Crick base pair but also the adjacent U10 and G26 are nicely paired (Figure 2.30b), one would never assume a shifted pK_a at A9N1. The adjacent unpaired bulge nucleotides A24 and C25 would be seem much more likely candidates. However, in solution there is no strong evidence for an A9-U27 base pair in AvD5 and ScD5 [281] and G26 is exposed to the solvent. In AvD5, we did not apply any base-pairing restraints for A9-U27 in the structure calculations. In the resulting structure ensemble U27 is never close enough to form hydrogen bonds to A9. On the other hand, C25O2 frequently — and sometimes also C25O2' — approach the amino group of A9 and the phosphate group of C25 is not too far away, providing electrostatic density and potential hydrogen bonding partners which could influence the pK_a at A9N1 (Figure 2.31a). Such approach of C25 to A9 is also observed in some instances of the recently more rigorously re-refined structures of ScD5 [279]. When G26 is in the minor groove conformation its N7 is close to A9N1, which would make a hydrogen bond to a protonated A9N1 possible (Figure 2.31b). There are thus possible ways in which a protonated A9 could be stabilized, even if unfortunately the resolution of the bulge conformation in our structure ensemble is not high enough to firmly establish the exact reason for the shifted pK_a at this site.

2.6.3 Biological relevance of the pH dependence in the bulge and in the catalytic triad

The cytosolic pH of the nitrogen fixating soil bacterium *A. vinelandii* is 8 [371] and also in the mitochondria of healthy *S. cerevisiae* cells — the habitat of the ai5γ group II intron — the pH of 7.5[372] lies above the pK_as of 6.4 and 6.7 in the bulge and catalytic triad, respectively, of AvD5. Even if glucose-starvation can lead to a considerable drop in pH in the cytosol and mitochondria of yeast cells[372], under normal circumstances in the intron's active state or rather in the process of D5 docking to D1, only a small fraction of A32 and even less of A9 are expected to be protonated. Nonetheless, the small fraction of molecules that adopt the low-pH-conformation at physiological pH might still play an important role by allowing the local structures to dynamically sample the conformations of the docked state. The HIV TAR hairpin has recently been shown to dynamically sample relevant ligand-bound conformations in its free form[46,47,373]. The perturbed pK_as might thus play a role in keeping bulge and catalytic triad flexible enough to be able to adapt to their tertiary binding partners in the process of intron assembly.

On the other hand, in the docked conformation of D5 both bulge and catalytic triad are engaged in multiple long-range interactions. In addition, the bulge also has undergone a significant conformational change (compare crystal and NMR structures in Figure 2.30), and so it is to be questioned if the observed pK_as in this region would be the same in the folded intron as in the isolated D5 domain in solution.

Acid-base equilibria help to keep the bulge flexible

In the bulge, it is especially hard to imagine that the pK_a of A9N1 would remain the same upon docking of D5, which involves the formation of a Watson-Crick A-U base pair. The pK_a is thus expected to play a role mainly in the folding process and not later on.

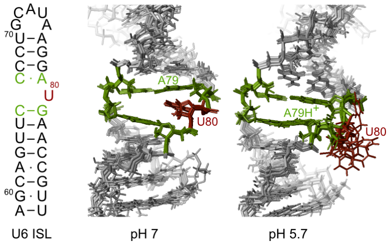


Figure 2.32 U6 ISL secondary structure and bulge structures at pH 7 (*left*, PDB-ID: 1SY4[374]) and pH 5.7 (*right*, PDB-ID: 1SYZ[298]).

A perturbed pK_a in the bulge region would, however, be interesting also because of a similarly shifted pK_a in snRNA U6, the proposed structural and functional homologue of D5 in the eukaryotic spliceosome (Section 1.4.1 in the Introduction and references therein). In U6, bulge nucleotide A79 is protonated with a pK_a of 6.5 (compare to the pK_a of 6.4 in the AvD5 bulge), in the course of which an A$^+$-C base pair is formed and the neighbouring unpaired U80 moves from inside the helix to a solvent-exposed conformation[297,298] (Figure 2.32). This makes a stacking of the two base pairs above and below U80 possible (green residues in Figure 2.32). The solvent-exposure of U80 when the pH is decreased bears some resemblance to the expulsion of bulge nucleotides A24 and C25 that happens when D5 converts to its docked conformation. In our AP24 fluorescence experiments we were aiming to answer the question whether A24 is also more solvent-exposed at low pH, i.e. more close to the docked conformation. Indeed, solvent exposure is increasing below pH 6.5, but the results remain ambiguous because also the effects of beginning acid denaturation of the helices can lead to a more solvent-exposed position of A24. In any case, the process is not governed by A9's pK_a of 6.4 but by a pK_a<6, which could reflect the protonation of A24N1 or C25N3. In summary, both in D5 and U6 ISL the elevated pK_as could be a necessary means to increase dynamics and thereby adaptability in the bulge and by that facilitate docking to the rest of the intron/spliceosome. It is furthermore possible that low pH biases the bulge towards the docked conformation.

Possible involvement of G4 flipping in group II intron catalysis

Neither A9N1 nor A32N1 are close enough to the scissile phosphate in the active site to suggest a direct involvement of their protonation/deprotonation in acid-base catalysis[106,236–238,284]. Nonetheless, the protonation in the G-A base pair in the catalytic triad of AvD5 might play a role in the catalytic mechanism by modifying the geometry of the active site.

Pairing of G4 via the Hoogsteen side in the protonated G-A$^+$ base pair leaves the Watson-Crick side available for interaction with a third nucleotide[20], which would encourage formation of a base triple. The G indeed forms a base triple with a G from the J23 linker in the only available crystal structures of a group IIC intron from *O. iheyensis* (G288 in Figure 2.33a, more generally designated G(J23) in the following text), but it does so in *anti*-conformation[236,238].

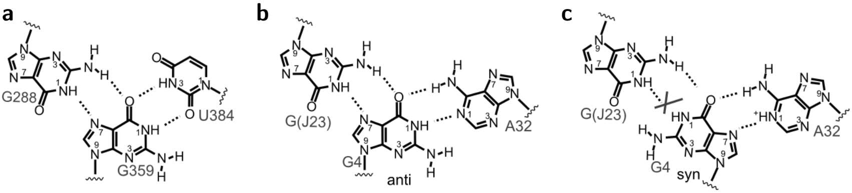


Figure 2.33 The base triple at the center of the catalytic triad. a) In the *O. iheyensis* group IIC intron crystal[106,236,238], G288 from the J23 linker forms a triple to the G-U wobble pair (G359-U384) of the catalytic triad. b) In AvD5 the G-A pair (G4-A32) at the same position in the catalytic triad can form a base triple in a similar way when G4 is in *anti*-conformation, but not in *syn*-conformation (c).

G(J23) is also conserved in Av.I5, in analogy we therefore expect also G4 in AvD5 in *anti*-conformation in a GA-G base triple (Figure 2.33b). Unless the orientation of G(J23) is different in AvD5, *syn*-conformation is not compatible with the base triple (Figure 2.33c). However, whereas none of the previous studies indicated significant conformational differences between the active site in the pre- and postcatalytic state or the substrate-bound and substrate-free state[236,238], the most recently published crystal structures[106] indicate a change in between the two steps of splicing. C25 (AvD5 numbering) and G(J23) are rotated and the interaction of the latter with the catalytic triad is broken[106] (PDB-ID: 4FAQ). In the resulting state also the catalytic metal ion cluster is disassembled. This intermediate state is suspected to be required to empty the active site and make it available for the substrate of the second splicing step. Similar substrate repositioning is also suspected in the spliceosome[375–377]). The flipping of G4 from an *anti*- to a *syn*-conformation interrupts the G4-G(J23) interaction and could in this way play a role in this rearrangement between the two splicing reactions (Figure 2.33c). In addition, the phosphate of G4, which is a ligand of the metal ion cluster, will be influenced by the switching of the G-A base pair and thus might help to disrupt the cluster.

The structure by Chan et al.[238] (PDB-ID: 4DS6) is thought to represent the pre-catalytic state because a G-to-A mutation at the position corresponding to G4 in AvD5 makes it inactive. Interestingly the introduced A was modelled in *syn*-conformation. The mutation leads to a slight modulation of the catalytic triad backbone and there is no conclusive electron density for the two catalytic metal ions and also the base of G288 (G(J23)) is missing. The structure is thus not only reminiscent of the *syn*-G4 in AvD5 but also resembles strongly the intermediate state in between the two splice steps proposed by Marcia and Pyle[106]. This might be a further clue for a physiological role of the *anti*- to *syn*-switch of G4.

The rearrangement in between the two splicing steps is also thought to go along with a conformational change of D6[106,246]. However, D6 is not observable in the *O. iheyensis* crystal structures, possibly because they lack the η-η' contact between D6 and D2. The intron only splices through the hydrolytic pathway, which uses water instead of the branch site 2'OH in D6[236]. In the absence of D6 we are still missing an important piece of the catalytic mechanism and it is not unlikely that the presence of the D6 branch site will introduce changes in the active site architecture, which might well reveal another role for the G-A mismatch and the *syn*-conformation of G4 in class E group II introns.

2.6.4 A strong Mg(II) binding site at the bulge

We have probed AvD5 with Mg(II) and its inner- and outer-sphere mimics Mn(II), Cd(II) and cobalt(III)hexammine, monitoring chemical shift changes and broadening, as well as intermolecular NOEs in the the case of cobalt(III)-hexammine in NMR spectra. The strongest binding is observed at the 5' phosphates and in addition to a slight preference for tandem G-C base pairs, both of which are already known in the literature[114,347]. Indication for another relatively strong Mg(II) binding site in the region of the bulge came from the biphasic behaviour of chemical shift perturbations when plotted against Mg(II) concentration. With the help of fluorescence experiments with the nucleotide analogue 2-aminopurine at position 24 of AvD5 a binding event with a K_D of about 0.27 mM consistent with the strong initial changes in the biphasic curves observed by NMR can be identified. However, no coordination to the bases of the bulge can be detected from the chemical shift perturbations of the ^{13}C and ^{15}N resonances (Appendix 14a). Taken together, we therefore suspect a Mg(II) ion binding to the phosphates in the bulge, which is also what happens in the fully folded intron[236]. Mg(II) binding increases the solvent exposure of A24, which could bias it towards the intron-docked conformation. The effect is dependent on pH and ionic strength, with a lower pH and lower ionic strength leading to stronger Mg(II) affinity.

2.6.5 The UAGUU pentaloop

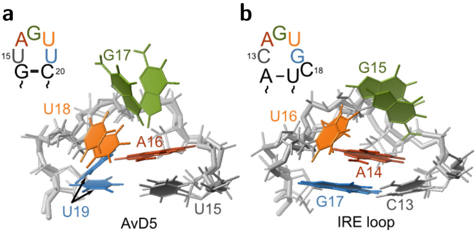


Figure 2.34 Comparison of the AvD5 pentaloop (a) and the loop of the IRE hairpin (PDB-ID: 1NBR[378]) (b). Both pictures show the two lowest energy structures of the respective NMR ensemble and their sequences are indicated on top. Corresponding residues in AvD5 and the IRE loop were colored alike. Arrows indicate the two orientations of U19 in AvD5.

The AvD5 pentaloop is not a conserved motif like the GNRA tetraloops often found at the same position in other group II introns. Even in the bacterial class E to which *A. vinelandii*'s intron 5 belongs, only the adenine at position 2 (A16) and the uracil at position 5 (U19) are conserved[235] (see Figure 1.15b in the Introduction). Motif searches to identify similar loops using RLooM[379] and AMIGOS II[380] are hampered both by the low conservation and the structural flexibility around nucleotides G17 and U18. Nonetheless, we find remarkable similarities with one published structure: the loop conformation of the IRE (Iron responsive element) hairpin, which was also studied in solution by NMR[378,381]. The 5'-CAGUGX-3' IRE loop is a pseudo-triloop, where C1 and G5 basepair and X is looped out (Figure 2.34b). The remaining central AGU triad is not only identical in sequence to the pentaloop in AvD5, but also shares the C3'-endo conformation of the G and the U (G17 and U18 in AvD5) and increased dynamics at these two residues[382,383]. Compared side-by-side, the two structures show different orientations of the guanine base (green in Figure 2.34a,b), but this base is expected to be flexible anyways (see Section 2.1.3). The orientation and position of all other loop residues is very similar. The G-C pair in the IRE loop is replaced by U15-U19 in AvD5, which is a less favourable base pair and our data consistently suggests that U19 is switching between U15 and A16 (Figure 2.34a).

There is also data on dynamic features of the IRE hairpin, which agree with the flexibility and base pair flipping we suspect at G17. ^{13}C relaxation studies of the IRE hairpin found lower order parameters in the loop than in stem purines at 37°C, but only a slight effect at 20°C[382]. Fluorescence studies employing 2-aminopurine in place of the adenine and guanine, respectively, of the loop indicate that the adenine is partly, the guanine mostly unstacked[383]. Concomitant stochastic dynamics simulations see the guanine sugar repucker on a nanosecond time scale and the base mostly in *syn*-conformation with short excursions to *anti*-conformation[383].

A crystal structure of IRE in complex with its protein binding partner IRP reveals that a distinct conformational change takes place upon binding. A and G bulge out, stacking onto each other and engaging in contacts with protein residues[384]. A similarly exhaustive rearrangement of the AvD5 pentaloop upon binding to a receptor structure is very plausible, not necessarily resulting in a bound conformation resembling the bound IRE loop. This behaviour is very different from the well described GNRA tetraloops, which do not undergo significant conformational change on binding to their respective receptor. It will be interesting to determine if the bound conformations of the AvD5 and IRE loops are similar, or if their adaptability allows for more varied binding modes.

2.6.6 Conclusion

Domain 5 has a central role in group II introns which is illustrated in a wealth of information that is already available around it (see Section 1.4.1 and references therein), not least the two solution and several crystal structures[106,236,238,280–282]. With AvD5 we not only add the first structure of a bacterial D5, with a significant variation in the catalytic triad region that allows a pH-dependent base flipping. We also add new evidence to the existing data on the well studied bulge region: 1) there is a strong pH dependence close to neutral pH at A9, 2) there is a metal ion binding site close to A24 with a much stronger affinity (K_D = 0.27 mM) than could be found previously. Finally, the newly solved structure of the capping UAGUU pentaloop in AvD5 reveals an unanticipated strong similarity with the IRE hairpin loop not only in terms of structure but also in terms of dynamics, suggesting a similar highly adaptive behaviour of this loop when docking to a tertiary interaction site.

3

The κ-ζ element of Sc.ai5γ

The κ-ζ element is of crucial importance for the initial compaction of D1 and folding of group II introns[258,320]. In later stages of folding it is the docking site of D5, interacting via both κ and ζ interaction sites. The element comprises nucleotides 195-218 and 374-386 of group II intron ai5γ from *S. cerevisiae* mitochondria. For NMR studies a 49-nt construct (D1kz) was chosen: G-C base pairs were added to the ends of helix d and d" for better transcription yield[12] and better stability and helix d" was closed by a UUCG loop (Figure 3.1). The construct is characterized by four helical regions (helix d, d', d" and the two base pairs in the κ extension), an 11-nt tetraloop receptor[271,385,386] (TLR, ζ) and a three-way junction which encompasses κ. In the following text we will use "κ" or "κ element" to refer to the four nucleotides A12, G21, A22 and A23 that make up the conserved motif, while "κ region" and "κ three-way junction" include also the neighbouring base pairs.

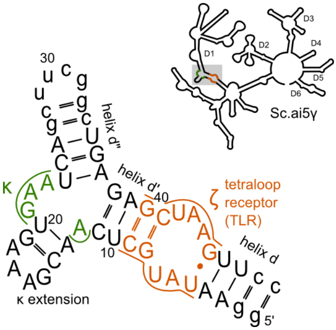


Figure 3.1 Secondary structure of the κ-ζ construct (D1kz) chosen for NMR studies. Small letters indicate nucleotides that are not present in the full intron structure. The small scheme on top indicates the location of the κ-ζ region in the secondary structure of group II intron Sc.ai5γ.

3.1 Screening for a suitable construct

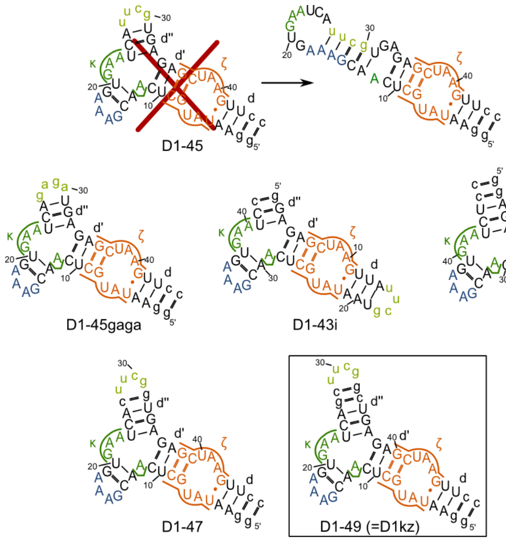


Figure 3.2 Constructs containing the κ-ζ region from Sc.ai5γ that were tested for their suitability for NMR studies. Small letters indicate nucleotides that were added to the wild-type sequence. The characterising structural elements κ, ζ, GAAA- and UUCG-loop are coloured consistently in all schemes. D1-45 was expected to form a three-way junction like the other constructs, but turned out to favour a hairpin conformation.

Before we arrived at the 49-nt D1kz construct (Figure 3.1) several variants containing the same κ and ζ parts but different numbers of added base pairs and different closing loops were screened to determine the most suitable candidate (Figure 3.2). The sequential walk region of their ^{1}H,^{1}H-NOESY spectra is shown in Figure 3.3. A first construct (D1-45) gave very sharp peaks but unfortunately the NOE connections showed that it favours a hairpin conformation over a three-way junction conformation. For all other constructs we found that the quality of NMR spectra depended strongly on the formation of helix d", which was not stable with only three base pairs even when closed by a loop (D1-43i, D1-45gaga). Five base pairs were still not enough when helix d" had open ends (D1-47i). Spectra of construct D1-47 with a four-base-pair d" helix closed by a UUCG loop looked much better. However, while the signature peaks of both the UUCG and the GAAA loop (known from the literature[281,387]) were immediately found in the spectra of D1-49 (=D1kz), they could not be identified in D1-47. Possibly D1-47 forms an alternative structure like D1-45. As D1-49 also resulted in resonably well resolved spectra, considering its size, this construct was selected for our further studies.

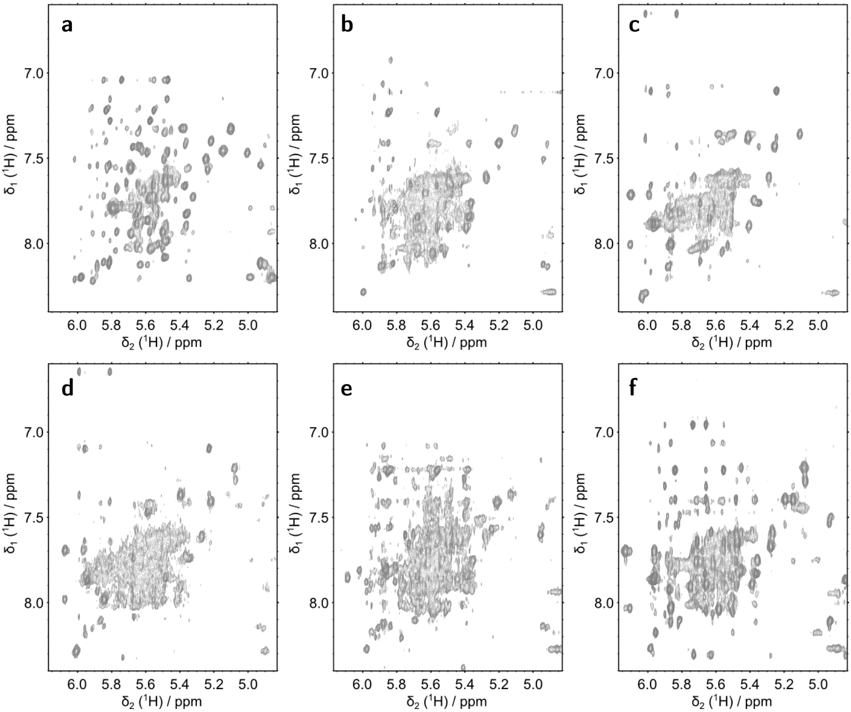


Figure 3.3 Sections of the ^{1}H,^{1}H-NOESY spectra of the screened κ-ζ constructs depicted in Figure 3.2. a) D1-45, b) D1-45gaga, c) D1-43i, d) D1-47i, e) D1-47, f) D1-49 (=D1kz). All spectra were taken at 297-300 K, 60 mM KCl, 10 µEDTA, pD 6.8, 600-700 MHz.

3.2 Spectral features and assignment in the absence of di- or trivalent metal ions

3.2.1 D1-27 – a helper construct for the assignment of D1kz

In addition to the RNA constructs described in the previous section we also prepared a shorter construct (D1-27, see inset in Figure 3.4a) to help in the assignment of the crowded spectra of D1kz. D1-27 contains ζ and helices d and d'. The latter is closed by a UUCG loop. Due to the much reduced number of residues compared to D1kz, the sequential walk could be followed straightforwardly and the assignments could be transferred well to the D1kz spectra (compare Figure 3.4a and b).

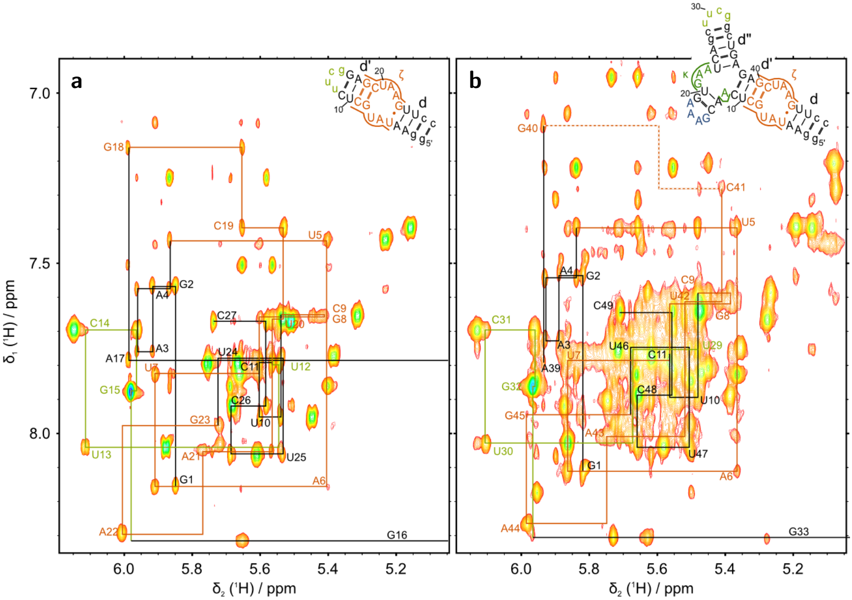


Figure 3.4 Sections of ^{1}H,^{1}H-NOESY spectra of a) D1-27 and b) D1kz. The sequential walk of D1-27 and the corresponding parts in D1kz (helices d and d' in black, ζ in orange and the UUCG loop in light green) are indicated. In b) the lines between G40 and C41 are dashed because the G40H1' resonance is unobservable (see also Section 3.2.2). G16H1'/G33H1' are strongly upfield shifted and not shown in this figure. Nucleotides that were added to the wild-type κ-ζ sequence are indicated by small letters in the secondary structure insets. (300 K, 60 mM KCl, 10 µEDTA, pD 6.8, 700 MHz)

3.2.2 Assignment of the sequential walk and other non-exchangeable protons in D1kz

The first NMR spectra of D1kz were recorded in the absence of di- or trivalent metal ions. The two tetraloops (the added UUCG and the GAAA in the κ extension) and the tetraloop receptor in D1kz are very common structural elements that have been studied previously in solution[281,385–387]. The published chemical shifts in all three regions compare well to the respective resonances in D1kz and were of great use in the assignment. Characteristic features of the UUCG loop like the strong intranucleotide H8-H1' peak of G32, which indicates *syn*-conformation, and the strongly shielded G33H1' and U29H2' are all present. In the GAAA tetraloop all characteristic resonances are observable except for G19H1', which is expected to resonate at about 3.6 ppm due to the ring current of the 5'-neighbouring nucleobase[357]. In this extreme position there is no spectral overlap that could hide the resonance, so we have to assume it to be broadened out by dynamics in the molecule and consequently not visible in our experimental conditions. In the tetraloop receptor (ζ), the sequential walk can be followed, but many resonances from A6 to G8 and from C41 to A43 (and G45) are broadened (Appendix 18a). Unexpectedly missing is the resonance of G40H1' situated inside the four-base-pair helix d'. This is all the more remarkable as the resonance is well visible in two constructs that are identical in this region, but lack the three-way junction (see Appendix 23 in the Appendix). G40H8-A39H1' and other A39 resonances are visible well enough, while the resonances directly at the three-way junction are severely broadened up to being unobservable. This affects in particular resonances from C11, A12, A13, A37 and G38 (Appendix 18a).

Sugar puckers and glycosidic angles. In 2D ^{1}H,^{1}H-TOCSY spectra the intensity of the H1'-H2' peaks indicates a C2'-endo conformation in all nucleotides for which it has been described in the literature (A6, A16, U30, C31 and A43)[385,387] and for none of the other residues. Similarly, no other nucleotide has an H8-H1' peak similarly intense as G32H8-H1' in ^{1}H,^{1}H-NOESY spectra, indicating that G32 is the only one in *syn*-conformation.

Since the majority of observable NOE cross peaks in RNA are of very short range (intranucleotide or sequential)[49], special interest arises for a group of medium- to long-range NOEs in D1kz which connects the adenines in κ to the minor groove of helix d'. Several of them (U10H1'/H2'-A22H2, A23H2-A12H1', A39H2-A22H1', A22H2-A39H2, A23H2-A39H1') are well observable in our experimental conditions in the absence of di- or trivalent ions. Others like A22H2-C11H1'/H2', A23H2-A12H2', C11H1'-A23H1', A12H8-A23H1'/H2' and A23H8-G21H2' can be assigned only in the spectra recorded in the presence of di- or trivalent ions. However, this is mostly due to resolved spectral overlap or sharpening of the involved resonances rather than to a substantial structural change, otherwise we could not expect to observe the first group of cross peaks either.

In summary, in the absence of di- or multivalent metal ions it is not possible to follow the sequential walk at several residues close to the κ three-way junction, which however does not prevent the observation of several NOEs which connect κ and helix d'.

3.2.3 Exchangeable protons and base pairing

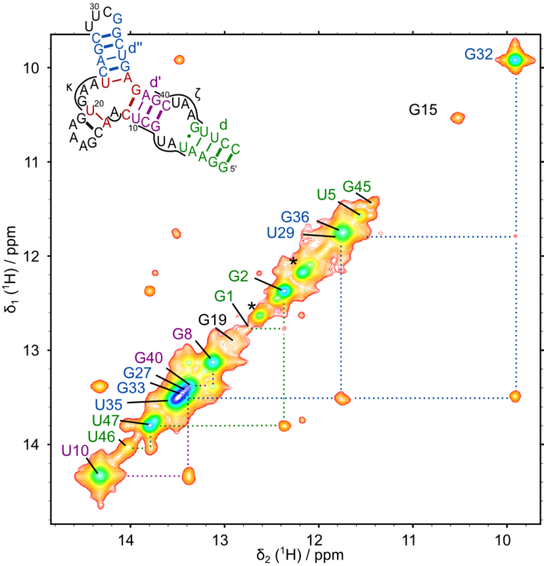


Figure 3.5 Assignment and connectivities of D1kz imino protons in a ^{1}H,^{1}H-NOESY section in the absence of Mg(II). Base pairs coloured red in the secondary structure inset are not observable in the spectrum. The asterisks indicate duplicate G1 and G2 imino resonances that are due to 3'-end heterogeneity (see Section 2.1.4) (0.65 mM D1kz, 285 K, 60 mM KCl, 10 µM EDTA, pH 6.75, 600 MHz).

In addition to the interruptions of the sequential walk, imino protons of the base pairs immediately adjacent to the three-way junction (C11-G38, A13-U20, U24-G36) are not observable in the absence of Mg(II) (Figure 3.5). Furthermore, dynamic processes at all neighbouring base pairs (U10-A39, C14-G19, C25-G36) are indicated by a distant broadening of the imino resonances when the temperature is lowered from 285 to 275 K. Sequential connections can be made for all observable imino protons, only G45H1-U46H3 and G15-G19 are often too weak to observe. J_{NN} HNN COSY experiments confirm in total only 9 of the 15 expected base pairs in the absence of Mg(II) (Appendix 24a).

Several imino resonances display a distinct temperature dependence (Figure 3.6). U5H3 and U45H3 at the tetraloop receptor as well as U29H3, an imino proton not involved in a base pair, broaden at higher temperatures, which could be a sign of an increased rate of exchange with the solvent. On the other hand, U10H3, G32H1, G36H1 and U46H3 are not visible in ^{1}H,^{15}N-HSQC spectra below 285 K, possibly indicating a fast conformational exchange that is pushed into the intermediate regime at lower temperatures (as shown in Figure 1.7a-c in the Introduction).

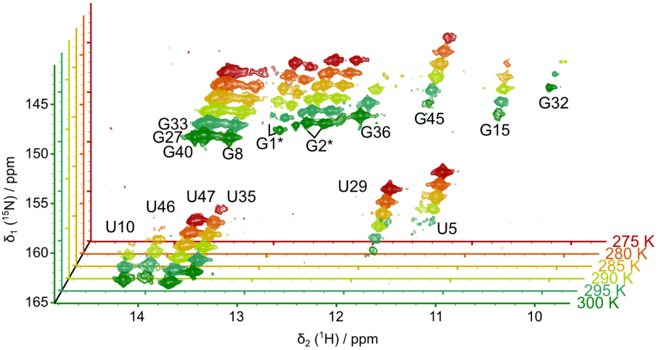


Figure 3.6 ^{1}H,^{15}N-HSQC signals of D1kz at different temperatures. The asterisks indicate duplicate G1 and G2 imino resonances that are due to 3'-end heterogeneity (see Section 2.1.4). (0.3 mM D1kz, 60 mM KCl, 10 µM EDTA, pH 6.8, 700 MHz)

3.3 The region of the κ-three-way junction is stabilized by di- and trivalent metal ions

3.3.1 Stabilization of base pairs neighbouring the three-way junction

Upon addition of Mg(II) several new imino peaks appear. They can be assigned to U20, U24 and G38, indicating the stabilization of all three base pairs directly adjacent to the three-way junction (Figure 3.7). A cross peak to U20 additionally allows the identification of the previously missing G21H1. Several cross peaks are observed to the new imino resonances, among them one connecting U24H3 to G38H1 across the three-way junction, which strongly indicates a coaxial stacking of helices d' and d" (red box in Figure 3.7).

Base pairs U10-A39, C14-G19 and C25-G36 are now observable in the J_{NN}

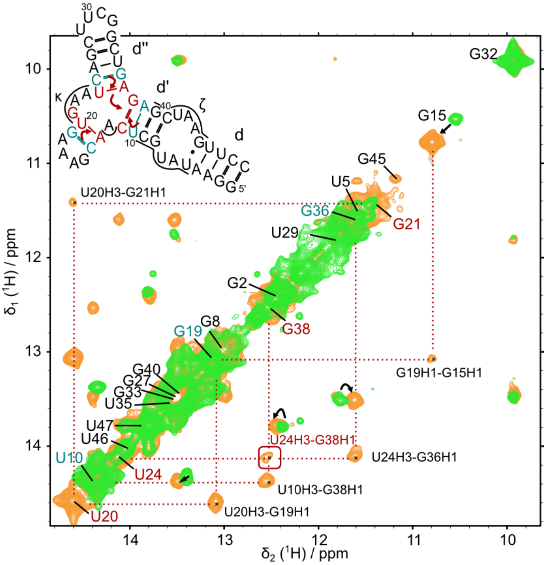


Figure 3.7 Newly appearing imino proton resonances in a ^{1}H,^{1}H-NOESY section in the presence of 10 mM Mg(II). The picture shows an overlay of the spectrum without (green) and with (orange) Mg(II). Cross peaks which appear with Mg(II) are labelled and the new NOE connectivities are indicated by red arrows in the secondary structure inset. Black arrows indicate peaks that move, but are present in both spectra. A red box indicates the cross peak which connects U24H3 and G38H1 across the three-way junction. (conditions (without Mg(II)/with Mg(II)): 0.65/0.6 mM D1kz, 285/282 K, 60 mM KCl, 10 µM EDTA, pH 6.75, 600/700 MHz)

HNN COSY (Appendix 24b), indicating that the exchange of their imino protons with the bulk solvent has been slowed down considerably. Also the base pair U24-A37 can be observed via both hydrogen bond donor and acceptor, which is still not possible in the other two base pairs directly adjacent to the three-way junction (A13-U20 and C11-G38).

The observed changes are not specific to Mg(II). The same resonances and cross peaks appear also in the presence of either Cd(II) or cobalt(III)hexammine (Appendix 25). However, both cobalt(III)hexammine and Cd(II) seem to be more efficient in stabilizing the region around κ than Mg(II). Not only do the imino protons in the region appear at lower ion concentrations (they are well observable already at 3.6 mM cobalt(III)hexammine and 3 mM Cd(II)), but peaks are also sharper and in particular U24H3-G38H1 — the most crucial NOE because it makes a connection from one side of the three-way junction to another — is better observable (Figure 3.7).

Monovalent salt alone on the other hand is not able to substitute for Mg(II). Apart from an increasing overall broadening and a slight shift of G8H1 no changes are distinguishable when KCl up to a concentration of 300 mM is added

(Figure 3.8).

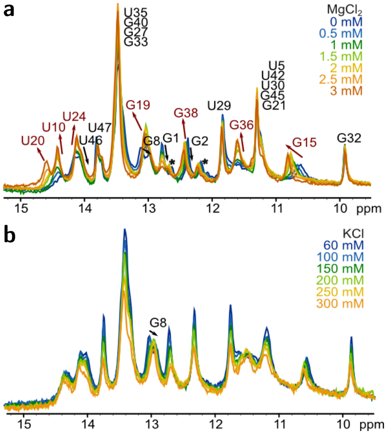


Figure 3.8 Imino region of 1D ^{1}H spectra in the presence of increasing amounts of MgCl$_2$ (a) and KCl (b), respectively. Resonances that grow sharper and more intense with Mg(II) are marked in red. Arrows mark the direction of the movement. The asterisks indicate duplicate G1 and G2 imino resonances that are due to 3'-end heterogeneity (see Section 2.1.4). Only G8 moves slightly in the KCl titration.

3.3.2 Newly appearing non-exchangeable proton resonances

When the Mg(II) concentration in an RNA sample is increased this usually leads to a general broadening due to unspecific metal ion–RNA interactions that happen at rates intermediate on the chemical shift time scale[142]. When titrating D1kz with Mg(II), however, we find that there are some resonances that become newly observable. While they can be discerned as very faint and broad signals already at 4 to 5 mM MgCl$_2$ when contour levels are set very low, they are visible more clearly at about 9 mM Mg(II) (∼15x the RNA concentration) (Figure 3.9). Unfortunately at 12 mM the spectra are already very much deteriorated by the broadening effect of unspecific Mg(II) interactions, and so no further improvement can be achieved.

The newly appearing peaks can be attributed mainly to resonances A12H2', A13H1', G19H1' and G38H1' (Figure 3.9). In addition, C11H1' and A37H1' sharpen distinctly (Appendix 18b). Resonances belonging to residues G21, A22 and A23 of the κ element and thus situated in the longest linker region of the three-way junction on the other hand do not change so dramatically. There are some chemical shift perturbations in κ (see Section 3.7.1), but the resonances do not sharpen or broaden remarkably and are observable equally well both in the absence and presence of metal ions. The most dramatic changes thus

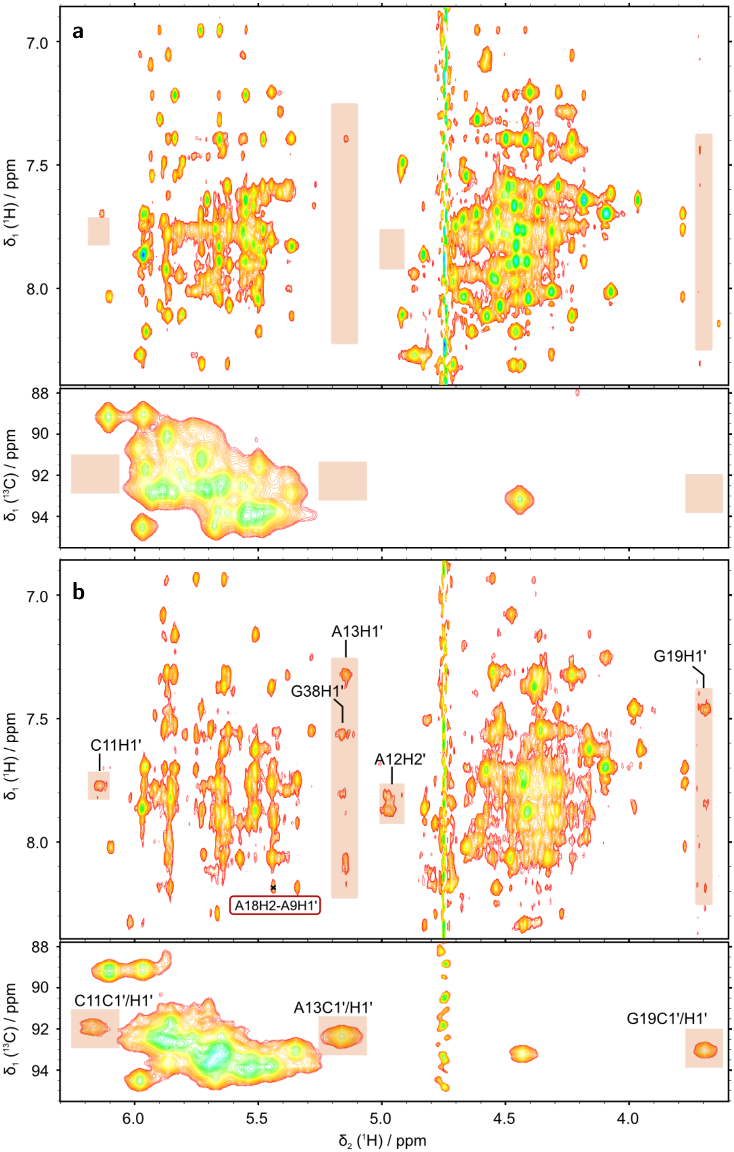


Figure 3.9 ^{1}H,^{1}H-NOESY and ^{1}H,^{13}C-HSQC sections in the absence (a) and presence (b) of 9 mM MgCl$_2$. Resonances that appear only in the presence of Mg(II) are highlighted in orange boxes. The intermolecular peak A18H2-A9H1', indicating a D1kz homodimer, is marked in a red box. (0.4-0.7 mM partially deuterated (NOESY) or ^{13}C,^{15}N-labelled (HSQC) D1kz, 300 K, pD 6.6-6.8, 60 mM KCl, 10 µM EDTA, 600 and 700 MHz)

do not concern the κ element itself but rather the nucleotides surrounding the three-way junction and the κ extension up to its GAAA tetraloop.

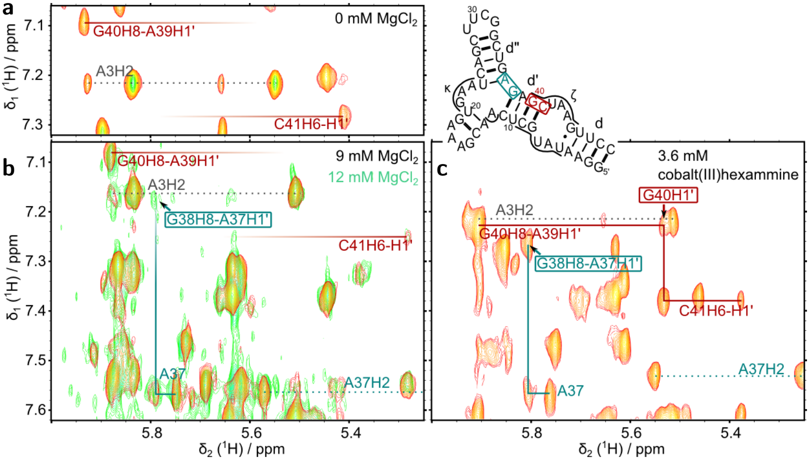


Figure 3.10 G40H1' and G38H8 are missing in the absence of metal ions (a) and in the presence of 9 mM Mg(II) (b). A very weak G38H8-A37H1' peak appears at 12 mM $MgCl_2$ (green spectrum in b). With 3.6 mM cobalt(III)hexammine (c) both G40H1' and G38H8 are well visible. The secondary structure scheme indicates the location of the respective nucleotides in the same colouring as the corresponding resonances in the spectra. The relatively invariable A3H2 resonance is indicated in all spectra as a point of reference. (0.35-0.7 mM D1kz, 300 K, pD 6.7-6.9, 60 mM KCl, 10 µM EDTA, 700 MHz)

Overall the same changes as with Mg(II) are also observed with cobalt(III)-hexammine (compare Figures 3.9 and Appendix 26) and Cd(II), but crucial peaks around the three-way junction are in many cases sharper and better observable. For instance, the newly appearing G38H8-A37H1' peak is not observable at 9 mM $MgCl_2$ and only vaguely perceptible at 12 mM $MgCl_2$ but clearly visible with 3.6 mM cobalt(III)hexammine (Figure 3.10). Furthermore, the G40H1' resonance that is absent in D1kz in all other tested conditions (without divalent ions, with Mg(II) or with Cd(II)) appears with cobalt(III)hexammine (Figure 3.10).

3.4 Dimer formation

By deciding to use the wild-type sequence of the κ-ζ region, we are working with a construct that contains both a GAAA tetraloop (TL) and its canonical

11-nt tetraloop receptor (TLR) sequence[271]. We did not expect significant problems due to intermolecular TL-TLR interaction because the affinity is known to be rather weak ($K_D = 0.4$ mM[388]) and interaction should not take place to a significant extent unless strengthened by additional interactions. In a previous study[386] the TL-TLR interaction could only be studied by NMR in solution in a molecule where TL and TLR were placed exactly one helix turn (11 base pairs) apart, thus making formation of a symmetric homodimer with two TL-TLR interactions possible. This substantially increases the affinity, putting it in the nanomolar range[386]. In D1kz it was not expected that the GAAA tetraloop and its receptor would be oriented in the required way with respect to each other to allow such a homo-dimerization.

Nonetheless, while in the absence of Mg(II) there is no sign of dimerization, in the presence of Mg(II) there is substantial evidence for the existence of D1kz dimers:

1) Native (non-denaturing) gels. While there is only one band for D1kz in the absence of $MgCl_2$ (upper band in lane two in Figure 3.11a), addition of 0.5 mM $MgCl_2$ to the running buffer causes a second band to appear higher up on the gel (lane 2 and 3 in Figure 3.11b). The relative band intensities of the D1kz monomer and dimer in Figure 3.11b suggest that about one fourth to one third of D1kz engages in dimers at this $MgCl_2$ concentration.

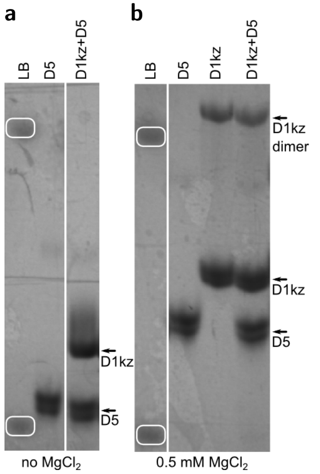


Figure 3.11 D1kz and D5 from Sc.ai5γ on native gels in the absence (a) and presence (b) of 0.5 mM $MgCl_2$. The $MgCl_2$ in (b) was added to the running buffer and gel solution. LB .. loading buffer containing xylene cyanol (upper band) and bromophenol blue as markers (lower band).

2) NOE cross peaks from the GAAA tetraloop to the tetraloop receptor. In the presence of 4 mM $MgCl_2$, a cross peak between A18H2 and C9H1' appears (Figures 3.9 and Appendix 18), together with a less pronounced cross peak between A18H2 and G8H2'. The A18H2-C9H1' cross peak appears also in the presence of cobalt(III)hexammine but not in the presence of Cd(II). These cross peaks would be expected for a cognate GAAA-tetraloop-receptor interaction, which would involve a base quadruple of G15-A18-G8-C41 and an AA platform involving A43 and A44[27]. Structure calculation trials in Xplor-NIH[348,349] make evident that neither A18H2-C9H1' nor A18H2-G8H2' can be satisfied by an intramolecular interaction without severely distorting the molecule.

3) Exchange cross peaks. When titrating D1kz with $MgCl_2$ we observe the appearance of new peaks along the diagonal in the aromatic-proton region of $^1H,^1H$-NOESY spectra (Figure 3.12). They connect resonances of the GAAA tetraloop with resonance lines that are otherwise completely empty (even the diagonal peaks are barely visible or absent). These characteristics are expected for cross peaks that are due to saturation transfer between a dominant main species and a very minor second species — thus the practically absent diagonal peaks — that occurs on the time scale of the mixing time, i.e. on the order of milliseconds[140].

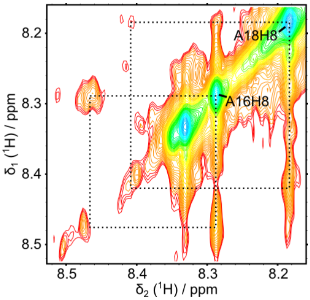


Figure 3.12 Exchange cross peaks of GAAA tetraloop resonances in a $^1H,^1H$-NOESY. (0.6 mM D1kz, 9 mM $MgCl_2$, 300 K, 700 MHz)

Such exchange cross peaks are observable for A16H8, A17H8, A18H8 and A16H2. Interestingly, the resonances of the indicated minor species agree very well with the published chemical shifts of a receptor-bound GAAA tetraloop (see Table 3.1), strongly suggesting that the observed cross peaks are due to exchange between its bound and unbound forms.

Table 3.1 Chemical shifts (in ppm) of GAAA tetraloop resonances in D1kz at 9 mM $MgCl_2$ compared to expected chemical shifts in the unbound and receptor-bound form of the tetraloop[281,386]. Residue numbering from D1kz is used. Resonances of the exchange peaks in D1kz agree very well with the ones found in the bound tetraloop.

	A16H8	A17H8	A18H8	A16H2
D1kz GAAA tetraloop[a]	8.29	7.92	8.18	7.75
D1kz GAAA tetraloop exchange peaks[a]	8.47	8.05	8.42	6.88
GAAA tetraloop unbound[b]	8.32	7.96	8.19	7.74
GAAA tetraloop bound[a]	8.51	8.01	8.33	6.96

[a] 9 mM Mg(II) (this work and Davis et al.[386]), [b] 0 mM Mg(II)[281].

Molecules that contain both a tetraloop and its receptor can hypothetically

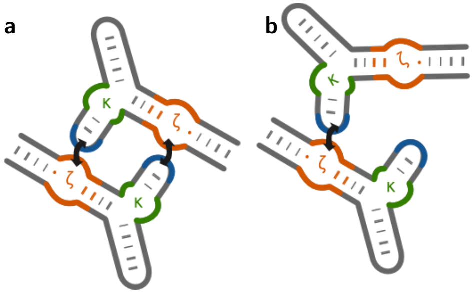


Figure 3.13 Schematic view of the two hypothetically possible homodimers of D1kz. While the symmetric homodimer in (a) has only one tetraloop-receptor interaction it is less likely to be sterically hindered than the symmetric homodimer in (b) that is suggested by our data. Orange: tetraloop receptor, blue: GAAA tetraloop, green: κ.

form two kinds of homodimers (Figure 3.13). Whereas a molecule has to meet certain conformational requirements to be able to form a symmetric dimer (Figure 3.13a), an asymmetric homodimer with only one tetraloop-receptor interaction (Figure 3.13b) is less likely to be sterically hindered.

In the classical TL-TLR interaction between the 11-nt TLR and a GAAA tetraloop we would expect strong characteristic chemical shift changes in several tetraloop and receptor proton resonances[386], which we do not observe here. On the other hand, we see exchange cross peaks between aromatic proton resonances of the GAAA tetraloop and resonances corresponding to the receptor-bound form of the tetraloop at 9 mM $MgCl_2$. This suggests that in contrast to the more stable contact between A18 and C9 (indicated by the sharp and pronounced A18H2-C9H1' peak), interaction of the rest of the tetraloop with the receptor is transient and only present in a very minor population of D1kz. In an asymmetric D1kz homodimer like in Figure 3.13b there is no steric reason, why the full canonical TL-TLR interaction should not form when there is a stable contact between C9 and A18 which is well compatible with it. In a symmetric dimer, on the other hand, a canonical TL-TLR interaction is likely to be prevented by steric hindrance but might be sampled transiently by one or both TL-TLR interaction sites. This would explain the exchange cross peaks between bound and unbound forms that we observe.

Taken together, our data suggests that D1kz forms a symmetric homodimer as shown in Figure 3.13a, where there is a stable interaction between C9 and A18, but the rest of the nucleotides which would be involved in a canonical TL-TLR interaction can only participate transiently in the interaction due to steric hindrance.

3.5 No signs of interaction between D1kz and D5 in solution

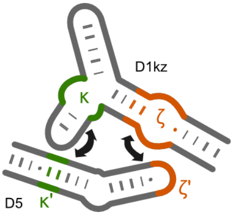


Figure 3.14 Schematic view of the conserved tertiary interaction sites κ and ζ of group II introns as present in D1kz and D5.

Upon group II intron folding, D5 docks into D1 interacting via the conserved κ-κ' and ζ-ζ' interaction sites[266,271,272] (Figure 3.14). In Sc.ai5γ both of them involve a GAAA tetraloop motif. ζ-ζ' is the canonical interaction of the D5 GAAA tetraloop with the 11-nt tetraloop receptor in D1[271]. The four nucleotides of κ have also been proposed to fold into a GAAA tetraloop motif[266], which has been confirmed more recently in the crystal structure of a group IIC intron[236,237]. The interaction partner of this GAAA motif is the minor groove of two G-C pairs in D5[236,266].

Docking of a GAAA tetraloop into the canonical 11-nt receptor leads to significant structural rearrangement of the receptor, new stacking and hydrogen bond interactions are formed. From previous NMR studies the characteristic NMR signature of these changes is known[386]. According to those, the biggest chemical shift changes are expected in D1kz for A43H1' (from 5.6 to 7.2 ppm) and A44H1' (from 5.9 to 3.7 ppm) in the TLR and in D5 for A16H2 (from 7.78 to 7.06 ppm) and A18H2 (from 8.15 to 8.81 ppm) in the GAAA tetraloop[386].

In several independent experiments we mixed D1kz and D5 and analysed NMR signals of both molecules. Experiments were done in the absence and in the presence of up to 12 mM $MgCl_2$ and with D1kz (0.25 to 0.6 mM) combined with D5 in ratios of 1:0.5 to 1:10. Before NMR measurements, the samples were heated to 65-70°C to separate the D1kz dimers and give D5 a chance to dock. In no case did we observe changes in the cross peak pattern of either D1kz or D5 that would indicate formation of a TL-TLR interaction. We also never observe exchange cross peaks involving the tetraloop resonances of D5 like we do for D1kz alone in the presence of high amounts of $MgCl_2$ (see previous section). This is in line with what we see on native gels, where no additional band is observed when D5 and D1kz are combined (rightmost lanes in Figure 3.11a and b).

Can the dimer formation be enough to prevent D1kz-D5 interaction in solution? Since we do not find any experimental evidence for an interaction between

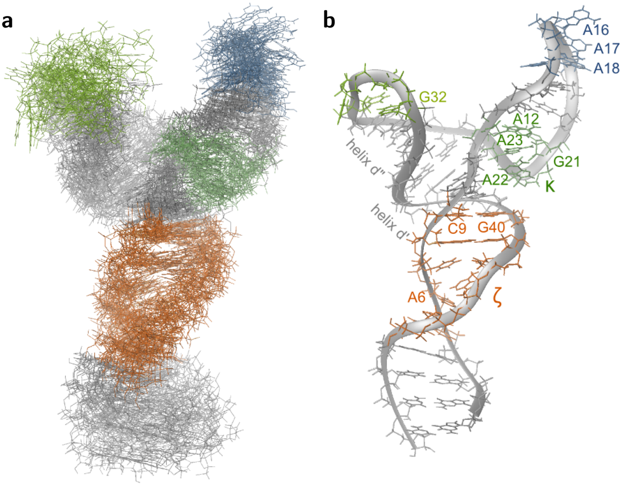
D1kz and D5 in the NMR studies and gel assays, but we do find it for the D1kz homodimer, the former has to be a less favourable interaction, in spite of the possible κ-κ' and ζ-ζ' contacts. The fact that there is no D1kz-D5 species observable when a majority of D1kz is still in its monomeric form (Figure 3.11) indicates that competition with the D1kz dimer is not the main reason why D5 does not dock.

3.6 D1kz structural features

3.6.1 Overall structure, RDCs

An ensemble of D1kz structures with an overall RMSD of 8.21 Å was calculated in restrained molecular dynamics simulations using 848 NOE distance restraints collected from spectra recorded in the absence and presence of 9 mM $MgCl_2$ (resonances peripheral from the junction are much sharper and allow for unequivocal assignment only in the absence of $MgCl_2$, while resonances near the κ three-way junction are only observable in the presence of $MgCl_2$). In case of experimental evidence for Watson-Crick base pairing from sequential imino proton cross peaks and J_{NN} HNN COSY spectra[58,389] and if the peak pattern in $^1H,^1H$-NOESY spectra agreed with A-helical conformation, hydrogen bonds and standard backbone angles (± 20°) were imposed. No backbone angles were restrained in the loop and bulge regions. Including a total of 35 residual dipolar coupling (RDC) restraints collected in the presence of $MgCl_2$ improved the overall RMSD to 4.3 ± 1.23 Å (Table 3.2). Both UUCG and GAAA tetraloop as well as the helical regions converge well (Table 3.2), with helix d' and d" almost forming a continuous helix (Figure 3.15). The three adenines of the tetraloop receptor mostly stack into the helix, which is typical for the undocked tetraloop receptor[385] but also other conformations with protruding A6 or A44 are compatible with the NOE restraints that are relatively weak in this region due to peaks broadened by dynamics.

In the smaller D1-27 construct (Section 3.2.1) more NOE cross peaks can be integrated well due to the reduced overlay and in general slightly sharper peaks, leading to a more well-defined tetraloop receptor region (Appendix 30).



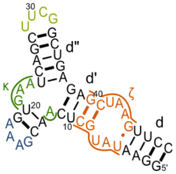


Figure 3.15 a) Ensemble of the 20 lowest energy structures of D1kz out of 200 calculated ones, aligned over all heavy atoms and b) the lowest energy structure alone. Orange: tetraloop receptor (ζ), light green: UUCG tetraloop, blue: GAAA tetraloop, dark green: κ.

Table 3.2 Structure calculation statistics for D1kz.

	without RDCs	with RDCs
RMSDs[a]		
overall (1-49)	8.21 ± 2.69	4.30 ± 1.23
κ extension (12-23)	2.24 ± 0.70	1.82 ± 0.45
helix d" (24-37)	1.56 ± 0.68	1.81 ± 0.66
κ extension and helix d" (12-37)	5.00 ± 2.10	3.24 ± 1.00
helix d, d' and d" (1-11,24-49)	7.39 ± 2.82	4.11 ± 1.35
κ, helix d' and d" (8-12,21-41)	3.44 ± 1.20	2.76 ± 0.83
distance restraints		
total	848	848
intranucleotide (i=j)	256	256
sequential (|i-j|=1)	420	420
medium (2≤|i-j|≤5)	38	38
long range (5<|i-j|)	134	134
dihedral restraints	276	276
H-bond restraints	83	83
RDC restraints	-	35

[a] Root mean square deviations (RMSDs) for the ensembles of 20 lowest energy structures calculated in Xplor-NIH are given.

3.6.2 The κ element forms an interrupted GAAA-tetraloop motif

The κ residues form a GAAA-like structure where A22, A23 and A12 stack (Figure 3.16a)[34,390]. Formation of a sheared G-A base pair between G21 and A12, as is typical for a GAAA structure, is compatible with our NOE restraints, but is not observed in all the structures. This is due to a scarceness of strong restraints in this region, owing to the fact that the sheared G-A base pair is not stable enough to allow the observation of its exchangeable protons. The H1' of the nucleotide following the GAAA tetraloop at the 3'-side (here A13), that is generally strongly shielded due to the ring current of the preceding adenine (see G19H1' which resonates at 3.79 ppm), is only moderately shielded in the present case. It resonates at 5.1 ppm, which is already a considerable shift for adenosine H1', which are usually found around 5.8 ppm. This could be indicative of dynamics in this region and an only transiently formed tetraloop structure or of a non-classical positioning of the nucleotide following the GAAA tetraloop.

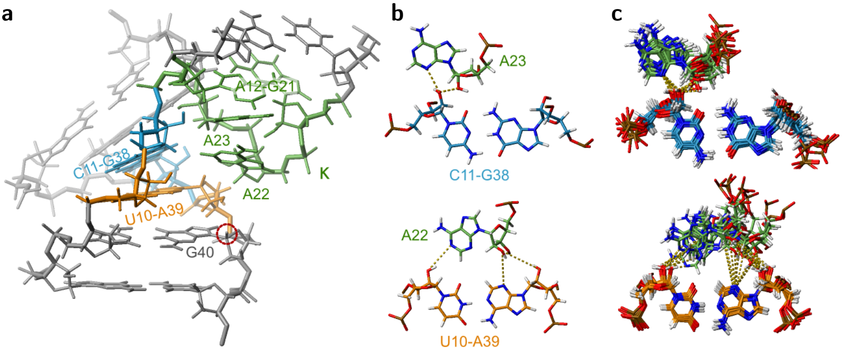


Figure 3.16 a) Details of the structure of the three-way junction and surrounding base pairs. κ and the base pairs that it forms A-minor interactions to are highlighted. A red circle marks G40H1', whose resonance is unobservable in all our spectra, probably due to a dynamic interaction with the adjacent κ. b) Top view of the two A-minor interactions in single structures and (c) in the ensemble. Suggested hydrogen bonds are drawn in yellow. The ensembles in (c) demonstrate the degree of uncertainty in the structures, which is caused by a scarcity of restraints due to dynamics in this region.

3.6.3 κ-residues A22 and A23 engage in A-minor interactions to helix d'

Several NOEs connect the adenines in κ to the minor groove of helix d' (see Section 3.2.2). The position of hydrogen bond donors at this interaction site in most of the calculated structures suggests the presence of hydrogen bonds from C11O2' to A23N3 and A23O2' as well as from U10O2' to A22N1 (and possibly from A22O2' to A39N3 and A39O2') in the context of A-minor interactions[25] (Figure 3.16b,c), even if the resolution especially around the A22-U10:A39 interaction is rather low (Figure 3.16c, lower panel) and the relevant distances usually exceed the length of a hydrogen bond. However, this is due more to the scarceness of information than to incompatibility: when the hydrogen bonds are included in the calculations as restraints, they are consistent with all experimental restraints.

The A-minor interactions can be classified as type 0 (A23N3 and O2' to C11O2') and type I (A22N1 to U10O2', A22O2' to A39N3 and O2') according to Nissen et al.[25] and follow the rule of decreasing in type order from 5' to 3'[25] (see Figure 1.4a in the Introduction).

The interaction of κ with helix d' also helps to understand why the H1' resonance of G40 is not observed in D1kz. At first sight this is unexpected considering that the residue is within a stable base pair in a helical region and that the H1' resonance is present and sharp in constructs lacking the κ region (see spectra of D1-27 and D1-45 in Appendix 23). The final structures of D1kz position G40H1' just below the contact interface of κ and helix d' (see Figure 3.16), where it transiently gets into contact with A22, inducing dynamics on a timescale that completely broadens the resonance.

3.6.4 The local structure in the three-way junction leads to a close approach of two backbone chains

Helices d' and d" form a continuous helix, where backbone strands leave and reenter between C11 and U24 (Figure 3.17a). Base pairs C11-G38 and U24-A37 stack onto each other, but there is also stacking between A12 and A23 in the κ motif. Together this forces the two backbone strands to cross in a very limited space leading to a region of concentrated negative electrostatic surface potential (Figure 3.17b). In particular, A12OP2 and U24OP1 are only ~4 Å apart in most of the calculated structures (Figure 3.17c), a configuration that is likely to require charge-compensation by metal ions. However, the fact that cobalt(III)hexammine is at least as efficient in stabilizing the local structure and the coaxial stacking of helices d' and d" as Mg(II), indicates that direct inner-sphere coordination is not required.

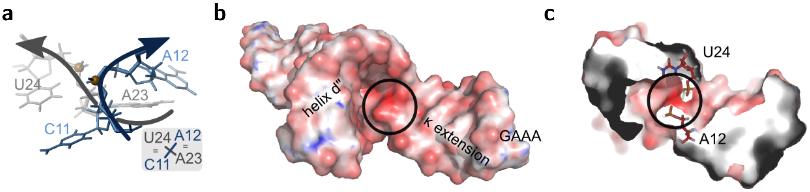


Figure 3.17 a) Detail of the junction region where the strands going to and coming from the κ extension cross. There is stacking between U24 and C11 and between A12 and A23, leading to a close approach of the phosphate backbones. b) At the crossing, there is a patch of higher negative electrostatic surface potential, which denotes the location of A12 and U24 phosphates (c). The potential is represented by a colour gradient from red (-1068 mV) to blue (134 mV).

3.7 Metal ion binding sites

Mg(II), Cd(II) and cobalt(III)hexammine exert a strong effect on D1kz in the region of the three-way junction and the κ extension (see Section 3.3). From the structural context (see previous section) one binding site is strongly suggested to be close to the phosphate groups of A12 and U24. To collect further evidence for metal ion binding in this and other regions of D1kz we looked more closely at the effects of the metal ions on resonances all over the molecule.

While the effect of Mg(II) on the ^{31}P resonances would be of especially strong interest, considering the indicated binding site at the phosphates of A12 and U24, it was not possible to attribute the ^{31}P signals due to the combination of low signal intensity and strong overlap. We therefore have to rely on the chemical shift changes of protons, nitrogens and carbons of the nucleobases which are more easily followed and can also indicate possible binding sites. The most direct information thereby is gained from the nitrogens, as they can directly coordinate the metal ions. By combining Mg(II) binding data with that of Cd(II), which is more likely to bind inner-spherely, and cobalt(III)hexammine, which can only bind outer-spherely, we aimed to get more insight into possible binding modes of Mg(II).

3.7.1 Mg(II)

Proton resonances of a partially deuterated sample of D1kz were attributed in ^{1}H,^{1}H-NOESY spectra and their chemical shifts plotted against increasing Mg(II) concentration. Some curves are clearly biphasic (Figure 3.18), indicating an overlay of a second effect which could be structural rearrangement or a second metal ion binding close-by. This behaviour mostly involves nucleotides located in the region of the GAAA loop (C14H6, G15H8, A16H8, A18H2', U20H6), κ (A22H1' and A22H2), the 3'- and 5'-ends (G1H8, C49H2') but also at some specific sites in helices (G36H2', A39H8, G40H8). The curves change their direction between 2 and 4 mM $MgCl_2$ (corresponding to 3-6x the RNA concentration).

In the strictly monophasic curves, chemical shift perturbations as a function of $MgCl_2$ concentration indicate K_D values in the millimolar range, comparable to the values found in the helical regions of AvD5, suggesting low affinity binding. Fitting the biphasic curves with an equation assuming two independent K_D values (equation 4.4) does not work well with the limited number of measurement points in the absence of additional knowledge about either of the two affinity constants involved. In the presence of multiple influences acting on the chemical shift, a quantitative analysis of affinity constants is therefore not feasible. However, qualitatively it can be noted that none of the proton chemical shift curves indicates a binding site with stronger than low millimolar affinity.

Figure 3.19 shows the absolute change in chemical shift experienced by sugar (H1' and H2') and aromatic protons (H2, H6 and H8) upon addition of 9 mM $MgCl_2$. The nucleotides that are most strongly influenced by the presence of $MgCl_2$ are C11 and A12 with resonances shifting by up to 0.4 ppm. C11 and A12 are in a critical position within the three-way junction, similar to A37 and G38, whose resonances also shift noticeably. In the GAAA tetraloop, perturbations are of average magnitude. Helix d" and the UUCG loop are remarkably unaffected by Mg(II), while the peaks of nucleotides located in the ζ region do shift considerably, get broader and eventually disappear when increasing the amount of metal ions. This could be due to metal ion binding, dimerization (see Section 3.4) or both.

To evaluate possible inner-sphere binding of Mg(II), ^2J-^{1}H,^{15}N-HSQC spectra

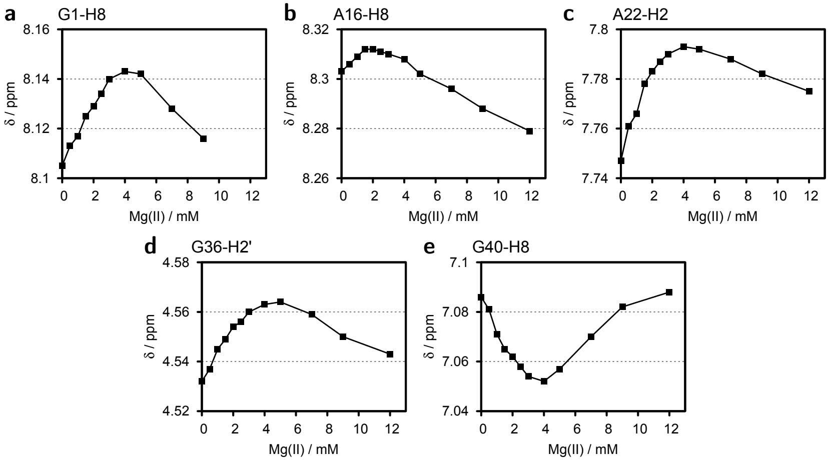


Figure 3.18 Chemical shifts in several regions of D1kz show a biphasic behaviour when plotted against increasing $MgCl_2$ concentration. a) 5'-end, b) GAAA loop, c) κ, d) helix d", e) helix d'.

were recorded. Thereby direct coordination of the metal ion to the nitrogen of guanines and adenines can be detected[391], the N7 of guanines being generally the most favourable binding site[91]. The chemical shift changes experienced by both N7/H8 in guanosines and adenosines as well as N1/N3/H2 in adenosines can be followed contemporaneously[391]. Usually, an upfield shift of the N7 resonance upon metal ion binding indicates a direct interaction[74,391,392] while the corresponding H8 and C8 are shifted downfield due to the electron-withdrawing effect of the coordinated metal ion. In the case of inner-sphere interaction a large ^{15}N chemical shift change is expected[74,147] (see Section 1.2.1 in the introduction). The rate of exchange in the first coordination shell of Mg(II) can be on an intermediate time scale[142,393], leading to a broadening and complete disappearance of the ^{15}N resonance.

Mg(II) as a hard metal ion prefers oxygen over nitrogen binding. If the local environment nonetheless allows a strong interaction it is expected to move the ^{15}N resonance upfield by 6 or even close to 20 ppm[74,146,147]. None of our observable N1, N3 or N7 move that much. G1N7 moves most strongly but still not more than 3.4 ppm at 4 mM $MgCl_2$ (8x the RNA concentration). Binding at the 5'-nucleotide is a general phenomenon due to the charge of the attached

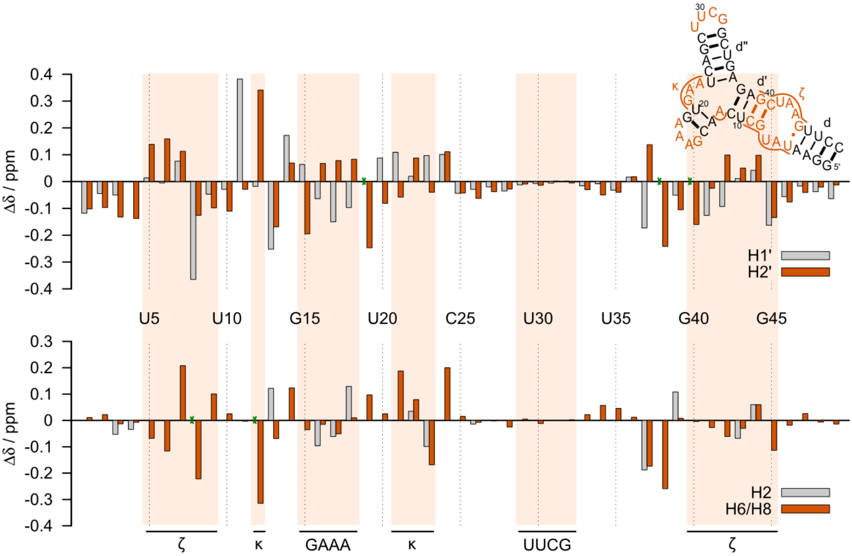


Figure 3.19 ^{1}H chemical shift changes in D1kz upon addition of 9 mM MgCl$_2$. To distinguish resonances that are not moving at all from values that are missing because assignment is not possible in all conditions the latter were crossed out green in the graph. (0.6 mM D1kz, 300 K, 60 mM KCl, 10 µM EDTA)

di- or triphosphate. Fitting the chemical shift changes of G1C8 and G1N7 with increasing Mg(II) concentration to a single binding isotherm yields an apparent K_D of 2.83 ± 0.15 mM. Since the apparent affinity is dependent on the affinity of other binding sites of the molecule it denotes a lower limit of the actual intrinsic affinity at this site[145].

Figure 3.20a shows the overlay of ^2J-^{1}H,^{15}N-HSQC spectra recorded with increasing amounts of MgCl$_2$. Residues in the helices (A3, A4, A26, G33, G36) show a moderate upfield shift of N7 (<1 ppm) and almost no shift of N1 and N3, which is consistent with a general non-specific interaction with the N7 and the bad accessibility of N1 and N3. In the ζ region N1 of A6, A43 and A44 and N7 of G45 broaden out and disappear upon addition of MgCl$_2$. This could suggest inner-sphere coordination in the tetraloop receptor region. On the other hand, also ^{1}H and ^{13}C resonances broaden out, and the cross peaks observed in ^{1}H,^{1}H-NOESY involving A43, A44 and G45 broaden more quickly in the presence of cobalt(III)hexammine where this effect cannot be due to

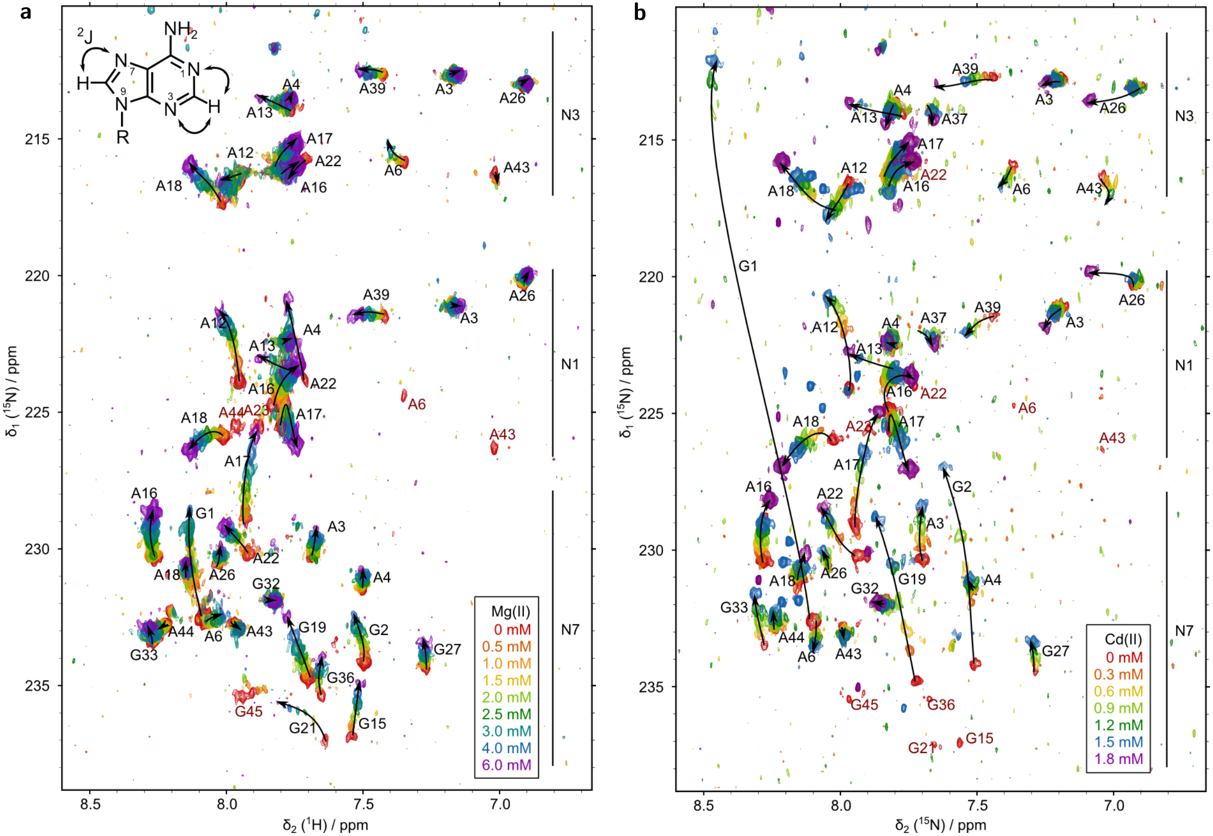


Figure 3.20 Overlay of ^{2}J-^{1}H,^{15}N-HSQC spectra of D1kz in the presence of increasing amounts of Mg(II) (a) and Cd(II) (b). Resonances that disappear at the first metal ion addition are marked in red. (0.5 mM D1kz, 300 K, 60 mM KCl and 700 MHz in (a), 60 mM KClO$_4$ and 600 MHz in (b))

inner-sphere coordination. In addition, we have described above that there are strong indications for a homodimer formation (see Section 3.4), which includes transient interactions of the GAAA tetraloop with the ζ region and can also lead to a broadening of the signals.

In the GAAA tetraloop and in the κ motif, the N7 tend to move upfield more strongly (>1 ppm) than in helix residues and without broadening. A17N1 and A18N1 turn around at about 2 mM $MgCl_2$ hinting at the interplay of structural rearrangements and metal ion interactions in this region (Figure 3.20a). In κ A12N1 and A22N1 move by almost 2 ppm, which is the strongest shift observed except for G1N7 and might be indicative of their good accessibility from the solvent.

The N7 of residues next to the three-way junction and in helix d' (G8, A12, A13, A23, A37, G38, A39, G40) are not observable in the absence or presence of metal ion. We can therefore not exclude strong binding to the nucleobase ^{15}N moieties at any of these sites. In the rest of the molecule, however, no strong, specific coordination of Mg(II) to purine nitrogens is observed.

3.7.2 Cd(II)

In contrast to Mg(II), the softer metal ion Cd(II) prefers the nitrogen ligands of RNA over the oxygen ones. In addition it has a less favourable solvation enthalpy and is thus more likely to interact in an inner-sphere fashion. Nonetheless, Cd(II) can coordinate at the same sites in the nucleobases as Mg(II), but will do so more tightly and thus potentially can yield sharper and better observable signals.

The changes observed in the proton chemical shifts of D1kz upon titration with $Cd(ClO_4)_2$ are very similar to those induced by Mg(II) and cobalt(III)-hexammine, large deviations are apparent only at G15H8, A26H2/H8, G33H8 and G36H8 (Appendix 31). Also ^{15}N shifts are similar to those induced by Mg(II), even if somewhat stronger: at 1.8 mM $Cd(ClO_4)_2$ the picture is already comparable to the one at 6 mM $MgCl_2$ (Figure 3.20), consistent with the stronger preference of Cd(II) for the nitrogen sites. In addition to the N7 of ζ-residues also G15N7, G21N7 and G36N7 disappear at the first addition of Cd(II), while they are observable throughout the Mg(II) titration. G15 and G21 occupy equivalent positions in GAAA-tetraloop motifs and it is reasonable that they should be affected in the same way, indicating that there is inner-sphere coordination of Cd(II) at this position. This is supported by comparatively strong downfield shifts of G15C8 and G21C8, which is expected in case of an inner-sphere coordination[74].

G36 reacts more strongly to Cd(II) than other residues in a helical context. This was already indicated by the proton chemical shifts and might be related to the major groove of G36 being faced towards the three-way junction, where a metal ion binding site is suggested by the negative electrostatic surface potential (Figure 3.21 and Section 3.6.4).

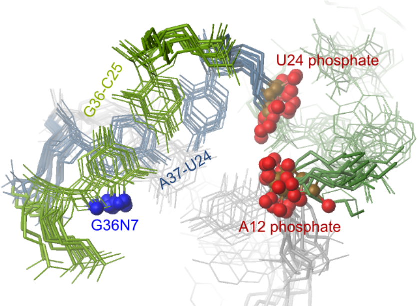


Figure 3.21 The suggested metal ion binding cavity close to the three-way junction of D1kz. Cobalt(III)hexammine and Cd(II) studies indicate an interaction with the nucleobase of G36, which faces into the same groove as the region of high electrostatic surface potential created by the phosphate groups of A12 and U24 (see Section 3.6.4). The ten lowest energy structures are shown in the picture.

However, binding at G15N7, G21N7 and G36N7 is not as tight as at the 5'-end, where G1N7 jumps by more than 20 ppm in the presence of Cd(II). The resonance disappears immediately after addition of 0.3 mM Cd(II) and reappears at 0.9 to 1.2 mM at this very upfield shifted position, which together with the 2.6 ppm downfield shift of G1C8 agrees well with what is observed in the literature upon direct coordination of Cd(II)[74]. Cd(II) in contrast to Mg(II) tends to form a macrochelate with N7 and the phosphates of a nucleotide[142,394]. The large chemical shift change of G1N7 indicates the formation of such a macrochelate in D1kz.

3.7.3 Cobalt(III)hexammine

D1kz was titrated with cobalt(III)hexammine up to 3.6 mM and observed by ^{1}H,^{1}H-NOESY spectra. Both the sign and magnitude of chemical shift perturbations induced by cobalt(III)hexammine are in general very similar to what we observe with Mg(II) (Appendix 31), indicating that the observed changes do not require inner-sphere coordination. The magnitude of changes at 3 mM cobalt(III)hexammine corresponds approximately to the one at 9 mM Mg(II).

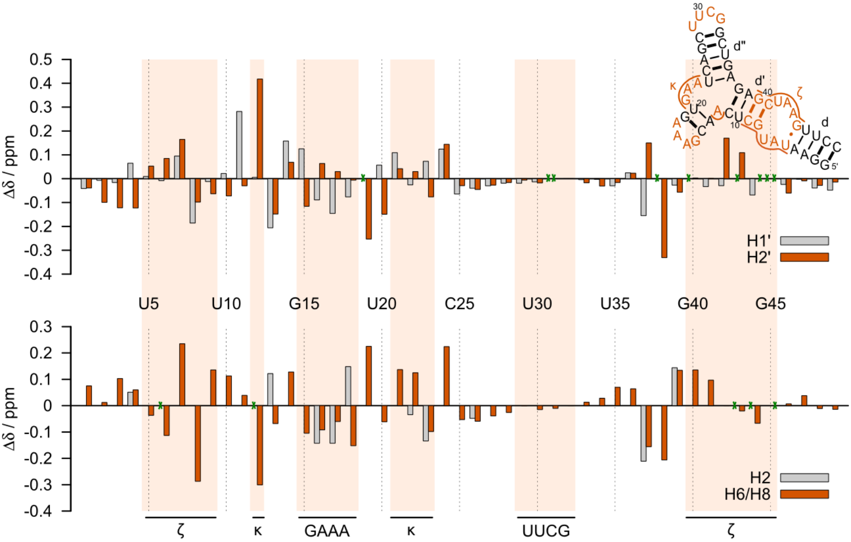


Figure 3.22 ^{1}H chemical shift changes upon addition of 3 mM cobalt(III)hexammine. To distinguish resonances that are not moving at all from values that are missing because assignment is not possible in all conditions the latter were crossed out green in the graph. (0.35 mM D1kz, 300 K, 60 mM KCl, 10 µM EDTA)

Only at the aromatic protons of GAAA tetraloop residues and of A39H8, G40H8 and C41H6, cobalt(III)hexammine seems to have a stronger impact than Mg(II) or Cd(II) (Appendix 31). The latter might be related to what we saw previously in this region: that cobalt(III)hexammine is also the only metal ion in whose presence the G40H1' resonance is visible, indicating that there is a specific stabilization.

At low concentrations of cobalt(III)hexammine (0.4 to 1 mM cobalt(III)hexammine, 1-2x the RNA concentration) we can identify three interaction sites in D1kz: (i) the GAAA tetraloop, (ii) the G-U wobble pair in the ζ region and (iii) the C25-G36/A26-U35 base pair step in helix d" (Figure 3.23 for non-exchangeable and Figure 3.24a for exchangeable protons). When the concentration is raised to 3.6 mM, cross peaks to the bulk resonance of cobalt(III)-hexammine appear to almost all imino resonances and there is strong overlap in the range of non-exchangeable proton resonances, indicative of diffuse ion binding. However, this higher amount of cobalt(III)hexammine is required in order to stabilize the resonances around the three-way junction and observe coordination in this region (Figure 3.24b).

Binding at G-U and GAAA tetraloop

The major groove of G-U wobble pairs provides an electrostatically favourable interaction site for cations[79], which could be the reason for cobalt(III)hexammine binding close to the U5-G45 wobble pair. Also the GAAA tetraloop has been observed to bind cobalt(III)hexammine already in a previous NMR study, where cobalt(III)hexammine was modelled into the major groove and makes close contacts to N7 of the guanine base of the closing sheared G-A mismatch as well as to phosphate oxygen atoms in the turn of the loop[80]. A K_D of 5±7 mM was determined from the chemical shift changes of the guanine imino proton. In our structure the cobalt(III)hexammine is located lower, more towards the N7 of A18 and G19 and further away from the loop phosphates and we determined a slightly lower apparent K_D of 1.79±0.69 mM from the chemical shift perturbations of non-exchangeable protons in the loop and closing base pair. We do not observe a cross peak to A16H8 like Rüdisser and Tinoco[80], which could be responsible for restraining the complex to this higher position in their structure. The different structural context as well as influence from the dimerization in D1kz can very well cause the observed differences in location and affinity of cobalt(III)hexammine.

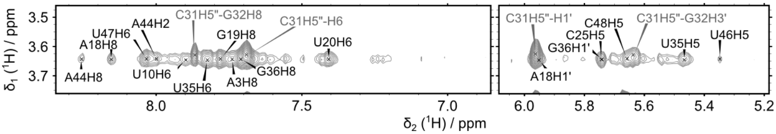


Figure 3.23 NOE cross peaks between aromatic and sugar protons of D1kz and the bulk cobalt(III)hexammine resonance at 3.7 ppm. (0.4 mM $Co(NH_3)_6Cl_3$, 0.35 mM D1kz, 300 K, 60 mM KCl, 10 µM EDTA).

Binding at the three-way junction

Even though at 3.6 mM cobalt(III)hexammine unspecific interactions obscure the more specific binding sites we need this higher concentration to observe crucial resonances around the three-way junction and determine if any cobalt(III)hexammine coordination happens in that region. In spectra with

3.6 mM cobalt(III)hexammine we therefore adjusted the contour levels in a way that only the most pronounced peaks would be visible (Figure 3.24b). While cross peaks to U47H3 and U5H3 are lost because the resonances broaden considerably, cross peaks to G15H1 and G19H1 are still the strongest signals. In addition, we also gain strong cross peaks to the neighbouring U20 and G21, which were previously unobservable. Weaker cross peaks are observed to imino resonances in helices d' and d" (U10H3, U35H3, G33H1, G36H1, G38H1). At such high metal complex concentrations (>10x the RNA concentration) those weaker cross peaks are more likely due to the low-specificity diffusive binding of ions in the major groove of purines (especially of guanines) than to a binding site that is specific to the D1kz three-way junction. Affinity for the κ extension is definitely stronger, both close to the GAAA loop and close to κ. There might be two similar binding sites close to each other at these two GNRA motifs, but they could also be covered by the rolling of one cobalt(III)hexammine along the major groove. Close to κ (U20H3, G21H1) the binding site could overlap with the one indicated by the electrostatic surface potential at the A12 and U24 phosphates (see Section 3.6.4).

Also the cross peaks close to G36 observed already at 1 mM cobalt(III)hexammine (Figure 3.23) could be related to the binding site at those phosphates. Metal ions attracted there are already close to G36, which is situated just on the opposite side of the groove between the κ extension and helix d" (the same is also suspected for Cd(II), see Figure 3.21).

Modelling of the cobalt(III)hexammine to three binding sites in D1kz

We used the NOEs collected at 0.4 to 1 mM cobalt(III)hexammine as distance restraints (6 Å to the central Co(III)) in restrained molecular dynamics simulations to position three cobalt(III)hexammine complexes in D1kz: at the G45-U5 wobble pair, at G36 and at the GAAA tetraloop (Figure 3.24c). It is possible that there is another binding site directly at the three-way junction, but we did not attempt to include it, because it is not possibly to unambiguously assign the observed NOE cross peaks in this region to only one binding site: each of the peaks around the three-way junction could also belong either to the binding site in the d" helix or to the one at the GAAA tetraloop. We can also not exclude that the cross peaks are due to a rolling of the cobalt(III)hexammine along the

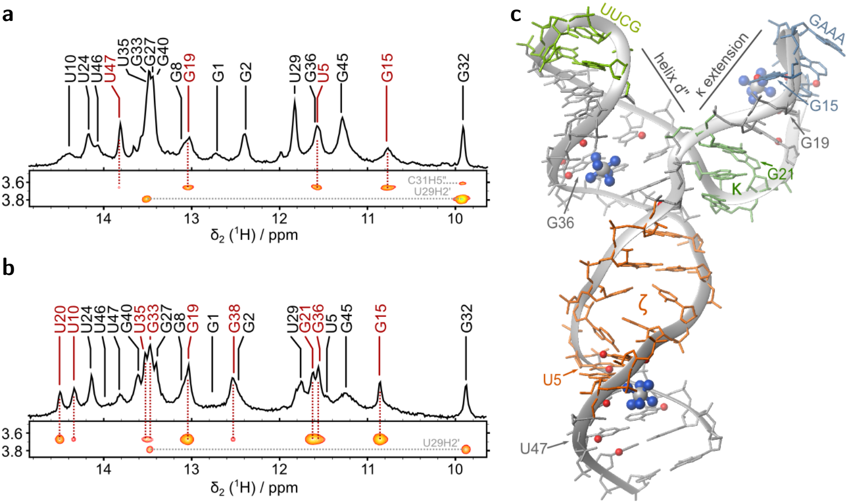


Figure 3.24 a, b) NOE cross peaks between D1kz imino protons and the bulk cobalt(III)-hexammine resonance at 3.7 ppm. 1D ^{1}H spectra and sections of ^{1}H,^{1}H-NOESY spectra in the presence of 1 mM (a) and 3.6 mM (b) $Co(NH_3)_6Cl_3$ are shown (0.4 mM D1kz, 275 K (a) and 278 K (b), 60 mM KCl, 10 µM EDTA, 700 MHz). c) Three cobalt(III)hexammine molecules were modelled into the D1kz structure on the basis of the NOE data collected at 0.4 to 1 mM cobalt(III)hexammine.

major grooves of the κ extension and helix d”.

3.8 D1kz - Discussion and conclusions

D1kz is not a static molecule. Apart from the tetraloop receptor region ζ, which has to be flexible to be able to rearrange between a free and a bound structure[385], the broadening or absence of resonances and lack of stable base pairs in the vicinity of the three-way junction indicate a flexible structure where metal ions do not bind to a preformed configuration but can participate in shaping the local structure. While main features of this structure can be discerned already in the absence of Mg(II) — a GAAA tetraloop motif of the four κ nucleotides and A-minor interactions to helix d' — the stable coaxial stacking of helices and close packing of backbone strands at the junction depend on di- or trivalent metal ions.

3.8.1 Metal ion binding sites

An electrostatic binding pocket at the three-way junction

NMR observables like chemical shifts and line broadening are affected both by changes in electron density due to a metal ion that is close-by and by conformational rearrangement. The overlay of the two effects makes it difficult — particularly in the region of the three-way junction — to define the binding site closely because signals in the whole region are strongly affected. Therefore the main evidence for a binding site at the phosphates of A12 and U24 comes from the structure rather than directly from the spectra. We see the two backbone strands pass each other quite closely, leading to a patch of negative electrostatic surface potential (Figure 3.17). The fact that most of the resonances that are broadened out in the absence of metal ions and appear upon their addition are situated close to this crossing further confirm that this binding site plays a critical role in the stabilization of the three-way junction.

No Mg(II) ion is found at this location in the crystal structure of an *O. iheyensis* intron and in a nucleotide analog interference mapping (NAIM) study on the D135 ribozyme a phosphorothioate at A12 only weakly interferes with tertiary collapse, at U24 not at all[258]. This and the fact that cobalt(III)hexammine as well as Cd(II) have a stabilizing effect equivalent to Mg(II) indicates that no specific coordination site is involved in the stabilization of the κ three-way junction, but rather a wider electrostatically favourable binding pocket situated between helix d” and the κ extension. This is supported also by signs of interaction of cobalt(III)hexammine and Cd(II) with the lower part of helix d” and the κ extension.

Differential binding at the GAAA tetraloop

Other binding sites are suggested at the GAAA tetraloop (Mg(II), Cd(II) and cobalt(III)hexammine) and at the G-U wobble pair at the base of the tetraloop receptor (cobalt(III)hexammine), both of which have already been reported in the literature[79,80,395]. The coordination site of cobalt(III)hexammine at the GAAA tetraloop is slightly shifted compared to previous solution and crystal structures[80,94] (see the positioning of a $Mg(H_2O)_6^{2+}$ in Figure 1.6b in the Introduction), which can be explained by the different structural context — vicinity to κ and the helix d” — but might also be related to the dimerization which takes place in the minor groove of the loop. It is thus not surprising that we find differences in the affinities compared to the published data, even if they are still in the low millimolar range indicating a rather weak interaction. Cobalt(III)-hexammine appears to bind a little more tightly in the GAAA tetraloop of D1kz than Mg(II), while Rüdisser and Tinoco[80] found it to bind a little more weakly (K_D=7±5 mM vs. K_D=2±1 mM for Mg(II) in 200 mM NaCl, 10°C, 10 mM sodium phosphate pH 5.6). This was attributed to an inner-sphere binding site of Mg(II) which is not accessible to cobalt(III)hexammine[80]. A localized inner-sphere coordination to the phosphates for Mg(II) was suggested also in a ^{31}P NMR study (K_D=2 mM in 100 mM NaCl)[395]. In this case also Cd(II) was investigated and found to bind with 10-fold stronger affinity, but in a slightly different manner with respect to G15 (G15P shifts differently). Our ^{15}N chemical shift data (Section 3.7.2) suggests that the difference might be an inner-sphere coordination of Cd(II) to G15N7 not present with Mg(II).

Direct Cd(II) coordination at the 5’-end

We observe the strongest ^{15}N chemical shift changes both with Mg(II) and with Cd(II) at G1N7. The first nucleotide contains a di- or triphosphate, which is an attractive binding site for both metal ions. It is known from studies on nucleotides[142,394] that simultaneous binding to the phosphates and inner-sphere coordination to N7 is favourable in the case of Cd(II), but not in the case of Mg(II), in line with what we observe at the 5'-end of D1kz. Upon addition of Cd(II) the G1N7 resonance disappears at first, but later reappears 20 ppm more upfield. This means that at low amount of Cd(II) the exchange between bound and unbound forms of G1N7 is in an intermediate regime on the chemical shift timescale which broadens out the signal. However, as the Cd(II) concentration increases the bound form becomes the most populated and can be observed in the spectra. For Mg(II), the affinity is much lower and the bound form of N7 can never be observed, even though the metal ion — possibly in its hexahydrated form — is attracted to the di- or triphosphate and thereby also exerts a shielding effect on the G1N7 resonance.

3.8.2 κ forms an A-minor junction in the solution structure of D1kz

Three-way junctions in published crystal structures have been classified into three families[396] on the basis of the relative lengths of the linkers connecting the three helices (Figure 3.25a). Family C is the most abundant one, found in various structured RNAs, and all its characteristics are found in the κ three-way junction: linker J31 (G21-A23 in D1kz, Figure 3.25a,b) is longer than J23 and J12 and helix P3 (the κ extension in D1kz) leans towards helix P1 (helix d"). In addition, the J31 linker can contact the shallow minor groove of helix P2 (helix d'), often in the form of A-minor interactions. Lescoute and Westhof[396] furthermore predict that the helices connected by a 0-nt linker (helix d' and d") will most likely be coaxially stacking, which is also true in D1kz in solution.

Apart from coaxial stacking and A-minor interactions another common building block of RNA structure is found in the small κ three-way junction: a GNRA tetraloop motif. It has been recognized long ago that this motif is a very favourable fold that can form even when its sequence is interrupted like in κ[34]. The κ three-way junction thus appears like a condensed example of structural principles of RNA and we were curious to find other examples in the available published RNA structures of the protein databank (PDB)[322]. However, database searches using FR3D[397] and RNA FRABASE[398] do not reveal a GNRA motif incorporated in a three-way junction in any other structure except the κ element of the *O. iheyensis* group II intron crystal structures[237], even less one where proximal A-minor interactions are formed like in D1kz in solution.

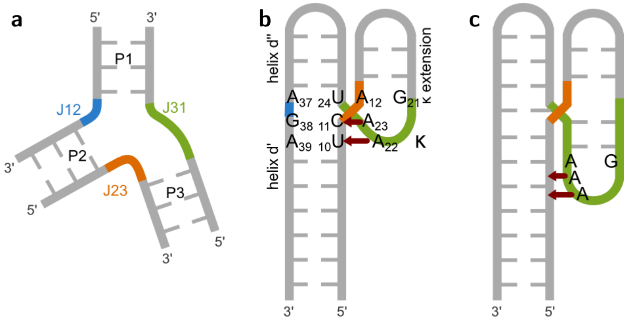


Figure 3.25 Classification of the D1kz three-way junction. a) Schematic view of a three-way junction with nomenclature used by Lescoute and Westhof[396]. b) Scheme of D1kz in the same colours as in a). c) Scheme of an A-minor junction in a four-way junction tectoRNA as described by Geary et al.[399].

However, in a way very similar constructs have been used by Geary et al.[399] to stabilize specific stacking conformers of RNA four-way junctions. They are termed “A-minor junctions” and based on the interaction of the terminal GNRA loop of one helix branch with the stem of another (Figure 3.25c). While in the four-way junction constructs there are some base pairs inserted between the junction and the GNRA tetraloop, in D1kz the crossing of the strands takes place directly in the motif, between the third and fourth nucleotide of the GNRA fold (compare Figure 3.25b and c). The three-way junction in D1kz thus represents the smallest A-minor junction possible, nonetheless capable of influencing the relative orientations of the helical arms with respect to each other.

However, while the A-minor interaction seems to determine the orientation of the κ extension in the solution structure of D1kz, in the crystal structure of a group IIC intron[105,236–238] other factors are more important. While κ forms a GNRA motif also in the fully assembled intron structure, it is oriented away from helix d’ and engages in H-bonds with the two G-C pairs of κ’ in D5 through the Hoogsteen face of A22 (N7), the W-C face of A23 (N1) and the sugar face of A12 (N3) (numbering of D1kz) (Figure 3.26).

Nonetheless, our spectra indicate that the A-minor interaction in D1kz is strong and independent of the presence of Mg(II). So what makes κ move away from helix d’, to a position which enables it to provide the second interaction site — in addition to ζ — for D5? It is possible that the docking of D5 itself switches

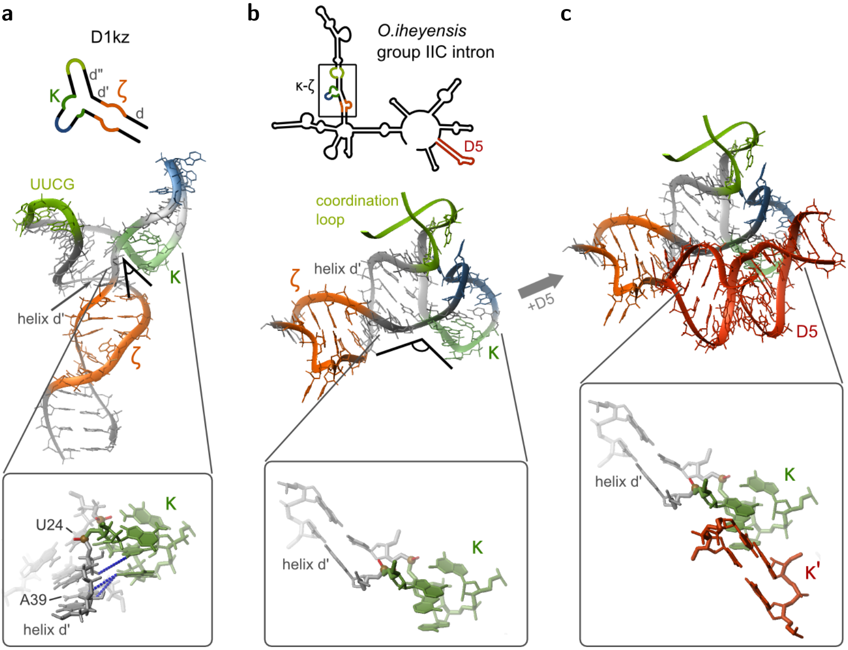


Figure 3.26 The orientation of the κ element differs significantly in D1kz and in the κ-ζ region of a fully assembled group II intron. a) In D1kz intramolecular A-minor interactions from κ-adenines bring κ close to helix d'. b) In the crystal structure of a group IIC intron from *O.iheyensis* (PDB-ID: 3G78[237]) κ is far away from helix d'. This is the orientation required for the docking of D5 (c). The panels at the bottom show details of the κ element. The phosphates of A12 and U24 and the equivalent atoms in the *O.iheyensis* intron are highlighted. Blue dashed lines in (a) indicate examples of observed NOEs. All depicted structures were superimposed on the heavy atoms of κ.

the κ interaction. However, we did not find any sign of D5 and D1kz interaction in solution (see Section 3.4), concluding that the interaction is too weak in the absence of additional tertiary interactions. Possibly the two constructs (D1kz and D5) only adopt global conformations that allow a favourable interaction when they are embedded in the complex framework of the folded group II intron. The weak affinity of D5 for D1kz in solution rather suggests that D5 will not be able to disrupt the A-minor interactions, even if we can not exclude that the situation is different inside the completely folded intron.

In the group IIC intron crystal structure[237,400] the GAACA capping loop of the κ extension (GAAA in D1kz) and the lower part of the neighbouring so-called coordination loop interact via stacking and H-bonds (Figure 3.26b). Even if this interaction is not abolutely required for a functional group II intron — when the GAAA tetraloop is replaced by UUCG in Sc.ai5γ, full catalytic activity is retained[240] — it could still be a contributing factor in making κ accessible to D5.

In the context of the full intron a combination of the two factors — the incoming D5 and the coordination loop — might control the final orientation of the κ motif.

3.8.3 Mg(II) promotes coaxial stacking of helices d' and d" in solution - the first step of group II intron folding?

Helices next to multi-branch junctions tend to maximize coaxial stacking to reduce overall folding free energy[401] and the stacking has been shown to depend on di- or multivalent ions[39].

In D1kz, stabilization of base pairs around the three-way junction and especially the appearance of an NOE cross peak between U24H3 and G38H1 are strong evidence for a stable coaxial stacking of helix d' and d" in the presence of Mg(II), Cd(II) or cobalt(III)hexammine. This coaxial stacking has been predicted by Costa et al.[277] from comparative sequence analysis and is observed also in the group IIC intron crystal structure[236,400]. The κ-ζ element has been shown to be of crucial imporance in the first step of splicing, the compaction of D1[258,320]. In particular a destabilization of helices d, d' and d" is noted in the absence of protein co-factor Mss116p or high Mg(II) concentration[258,320,321].

We could now show in D1kz that the stacking and a concomitant stabilization and rigidification of the adjacent helices is metal-ion dependent, but independent of the presence of tertiary interactions from the rest of the intron. D1kz can thus serve as a model for the earliest steps of group II intron folding, where tertiary contacts are not yet formed, and might explain at least part of the metal ion dependence of the process.

4

Experimental Section

4.1 Materials and chemicals

All chemicals were purchased from either Fluka-Sigma-Aldrich (Buchs, Switzerland) or Brunschwig Chemie (Basel, Switzerland) at puriss p.a. or biograde. All solutions were prepared using autoclaved and filtered ultrapure water. Aqueous acrylamide solution (AccuGel 29:1) and TBE (Tris-borate-EDTA) buffer for the purification of RNA and DNA were purchased from National Diagnostics (Hussle Hull, UK). The deuterated reagents D_2O (100% or 99.98% D), NaOD (99.9% D, 1 M in D_2O), DCl (99.5% D, 1 M in D_2O), were from Armar Chemicals (Döttingen, Switzerland). Buffer solutions (pH 4.00, 7.00 and 9.00) for electrode calibration were purchased from Metrohm AG (Herisau, Switzerland).

Nucleoside 5'-triphosphates (NTPs) were obtained from GE Healthcare (Glattbrugg, Switzerland), with the exception of UTP that was purchased from Acros-Organics (Geel, Belgium). ^{13}C and ^{15}N labelled NTPs were purchased from Silantes GmbH (München, Germany) while partially deuterated NTPs (ribose-[3',4',5',5"]- and pyrimidine-[5]-D) were obtained from Cambridge Isotope Laboratories (Andover, USA).

DNA oligonucleotides, with and without 2'-methoxy modifications at the two 5'-nucleotides, were purchased from Microsynth (Balgach, Switzerland). They were purified by polyacrylamide gel electrophoresis (PAGE) before usage, when they were not ordered already PAGE-purified. An RNA oligonucleotide containing a 2-aminopurine nucleotide was obtained from IBA (Göttingen, Germany).

T7 polymerase for the *in vitro* transcription of unmodified RNA samples was produced in-house[402].

Pf1 filamentous bacteriophage (RNase and protease free, in 10 mM potassium phosphate buffer pH 7.6) was purchased from ASLA Biotech Ltd (Riga, Latvia).

Cobalt(III)hexammine was synthesized in-house according to published protocols[403]. $MgCl_2$, $MnCl_2$ and $Cd(ClO_4)_2$ stock solutions for the metal ion titrations were obtained as ultrapure solutions (1 M in H_2O) from Fluka-Sigma-Aldrich (Buchs, Switzerland). The exact concentration of the $MgCl_2$, $MnCl_2$ and $Cd(ClO_4)_2$ stock solutions in 100% D_2O was determined by potentiometric pH titration with EDTA[404], while the one of $Co(NH_3)_6^{3+}$ was measured by UV-Vis[405] ($\epsilon_{476nm} = 56.5\ M^{-1}\ cm^{-1}$ and $\epsilon_{340nm} = 46.1\ M^{-1}\ cm^{-1}$).

4.1.1 Buffers

Urea loading buffer for denaturing PAGE. 11.8 M urea, 8.3% sucrose, or 4.2 mM Tris-HCl (pH 7.5), 0.8 mM EDTA (pH 8.0), 0.08% xylene cyanol (XC) and 0.08% bromphenol blue (BB)

Formamide loading buffer for native PAGE. 10 mM EDTA (pH 8.0), 0.16% xylene cyanol (XC) and 0.16% bromphenol blue (BB), 82% formamide

TBE buffer for denaturing PAGE and electroelution (10x). 0.89 M Tris-HCl (pH 8.3), 0.89 M boric acid, 0.002 M EDTA

Native PAGE buffer (5x). 0.33 M HEPES, 0.17 M Tris, pH 7.4.

Transcription buffer (5x). 200 mM Tris-HCl (pH 7.5), 200 mM DTT, 10 mM spermidine.

4.1.2 Instrumentation

The H_2O was treated with a TKA genepure water purification system from TKA Wasseraufbereitungssysteme (Niederelbert, Germany) prior to every experiment.

The electroelution apparatus used to recover the RNA was a Whatman Elutrap System from Whatman (London, UK) with BT1 and BT2 membranes. For desalting, either Vivaspin 2-mL ultrafiltration devices (Sartorius Stedim biotech, Aubagne, France) with 3 kDa (for D1kz also 5 kDa) cutoff membranes or NAP-10 columns from GE Healthcare (Glattbrugg, Zürich) were used. Centrifuges were a 5415R, 5415D and 5804R with rotor F-45-24-11 and 5804R with rotor A-4-44 from Eppendorf (Hamburg, Germany) as well as a RC5C plus from Sorvall (Langenselbold, Germany) with a SA-600 rotor. The purified RNA samples were vacuum-dried in a Concentrator 5301 from Eppendorf (Hamburg, Germany).

pH values were measured using a Hamilton Minitrode glass electrode from Hamilton AG (Bonaduz, Switzerland) connected to a Metrohm 605 digital pH-meter. Potentiometric pH titration were carried out with an E536 potentiograph connected to an E665 dosimat and a 6.0253.100 Aquatrode-Plus combined macro glass electrode from Metrohm AG (Herisau, Switzerland).

NMR spectra were recorded on a Bruker Avance 700 MHz spectrometer equipped with a CP-TXI z-axis pulsed field gradient CryoProbe, on a Bruker Avance 600 MHz equipped with a TCI z-gradient CryoProbe, and on a Bruker Avance 500 MHz, equipped with a QNP z-gradient CryoProbe. In the context of the Bio-NMR project, additional spectra were acquired on a Bruker Avance 900 MHz equipped with a TCI z-gradient CryoProbe at CERM (Centro di Ricerca di Risonanze Magnetiche) in Florence, Italy.

UV (Ultraviolet) spectroscopic measurements were carried out on Cary 100 UV-Vis and Cary 500 scan UV-Vis-NIR spectrophotometers from Varian Inc. (Palo Alto, USA) connected to a circulating temperature controller using 10 mm or 1 mm QS cuvettes from Hellma (Müllheim, Germany). For the determination of RNA concentrations also a NanoDrop 2000 spectrometer from Thermo Scientific (Waltham MA, USA) was employed. Fluorescence was measured on a Perkin-Elmer luminescence spectrometer LB-50 using 3 mm square cuvettes from Hellma (Müllheim, Germany).

4.1.3 RNA constructs

AvD5: 5'-GGAGCCGUAUGCGGUAGUUCCGCACGUACGGAUCU-3' (M_w = 11.3 kDa, $\epsilon_{260nm} = 379.8$ mM^{-1} cm^{-1})

AvD5-U35C: 5'-GGAGCCGUAUGCGGUAGUUCCGCACGUACGGAUCC-3' (M_w = 11.3 kDa, $\epsilon_{260nm} = 377.3$ mM^{-1} cm^{-1})

AvD5-AP24: 5'-GGAGCCGUAUGCGGUAGUUCCGC(AP)CGUACGGAUCC-3' (M_w = 11.3 kDa, $\epsilon_{260nm} = 315$ mM^{-1} cm^{-1}) (AP ... 2-aminopurine)

ScD5: 5'-GGAGCCGUAUGCGAUGAAAGUCGCACGUACGGUUCC-3' ($M_w = 11.7$ kDa, $\epsilon_{260nm} = 352.9$ mM^{-1} cm^{-1})
D1kz: 5'-GGAAUAUGCUCAACGAAAGUGAAUCAGCUUCGGCUGAGAG-CUAAGUUCC-3' ($M_w = 15.8$ kDa, $\epsilon_{260nm} = 494.7$ mM^{-1} cm^{-1})
D1-27: 5'-GGAAUAUGCUCUUCGGAGCUAAGUUCC-3' ($M_w = 8.6$ kDa, $\epsilon_{260nm} = 266$ mM^{-1} cm^{-1})
D1-45: 5'-GGAAUAUGCUCAACGAAAGUGAAUCAUUCGUGAGAGCUA-AGUUCC-3' ($M_w = 14.4$ kDa, $\epsilon_{260nm} = 463.5$ mM^{-1} cm^{-1})
D1-45gaga: 5'-GGAAUAUGCUCAACGAAAGUGAAUCAGAGAUGAGAG-CUAAGUUCC-3' ($M_w = 14.5$ kDa, $\epsilon_{260nm} = 472.8.5$ mM^{-1} cm^{-1})
D1-47: 5'-GGAAUAUGCUCAACGAAAGUGAAUCACUUCGGUGAGAGC-UAAGUUCC-3' ($M_w = 15.1$ kDa, $\epsilon_{260nm} = 479.6$ mM^{-1} cm^{-1})
D1-43i: 5'-GGAGAGCUAAGUUAUUCGUAAUAUGCUCAACGAAAGUG-AAUCC-3' ($M_w = 13.8$ kDa, $\epsilon_{260nm} = 448.0$ mM^{-1} cm^{-1})
D1-47i: 5'-GGAGAGAGCUAAGUUAUUCGUAAUAUGCUCAACGAAAG-UGAAUCUCC-3' ($M_w = 15.1$ kDa, $\epsilon_{260nm} = 487.6$ mM^{-1} cm^{-1})

4.2 RNA sample preparation and analysis

RNA samples were transcribed and purified following standard procedures[402]. The reaction mixture contained 5 mM of each NTP, 0.5-1.2 µM of the double stranded DNA template, 0.1% Triton X-100, 40 mM Tris-HCl (pH 7.5), 40 mM DTT, 2 mM spermidine and 30 mM (D1kz) or 35 mM (AvD5) $MgCl_2$. Typically 5 to 10 mL transcriptions were done, yielding 2-3 NMR samples (0.5-1 mM). The amount of polymerase depended on the effectiveness of the polymerase itself and was optimized for each polymerase batch. In order to avoid 3'-end inhomogeneity due to run-off transcription by the T7 RNA polymerase we used a modified DNA template, in which the two nucleotides at the 5'-end contain a 2'-methoxy group[328,329]. Transcriptions were carried out for 4 h or overnight in 1 mL aliquots in Eppendorf tubes, gently shaking at 37°C and followed by ethanol precipitation to remove excess salt. RNA was purified overnight at room temperature on denaturing polyacrylamide gels (15-18% acrylamide, 7 M urea, 1x TBE; 10-20 W per gel, size: 43x30x1 cm). RNA bands were located on the gel by UV shadowing at 254 nm and excised, followed by electroelution.

RNA was desalted by ultrafiltration in several washing steps, first with 1 M KCl, pH 8, then with ultrapure water. KCl was employed in particular to remove excessive Tris stemming from the TBE buffer used for purification, which can compete with RNA for metal ion binding[406]. The samples were lyophilized and redissolved in 250 to 300 µL 100% D_2O or H_2O/D_2O (9:1) usually containing 60 mM KCl, 10 µM EDTA and at pH 6.5-6.9 for all constructs (D1kz, AvD5 and ScD5). RNA concentrations were determined from the absorbance at 260 nm.

The pH in the samples was adjusted using 0.1 M DCl and 0.1 M NaOD solutions. In the case of samples prepared in 100% D_2O, 0.4 pH units were added to the pH meter reading to determine the pD[407].

For analytical purposes small non-denaturing (native) gels (11% acrylamide in 1x native gel buffer, gel size: 25x17 cm) were prepared without and with 0.5 to 30 mM magnesium acetate and run for 4-8 h at 5-10 W per gel.

4.3 NMR spectroscopy

RNA samples were transferred into 5 mm Shigemi tubes for all NMR experiments. ^{1}H chemical shifts were directly referred to external DSS (0.2%, pH 7.5), ^{13}C and ^{15}N were indirectly referred to ^{1}H of DSS[408].

Non-exchangeable protons were assigned from $^{1}H,^{1}H$-NOESY spectra in D_2O, recorded at various temperatures (290, 295, 300, 305, 310 K) and with different mixing times (60, 150, 250, 350 ms) on samples prepared with partially deuterated or natural abundance isotope nucleotides. Water suppression was either achieved by employing a water presaturation pulse or an excitation sculpting pulse scheme. $^{1}H,^{1}H$-TOCSY spectra (50 ms mixing time) were acquired to qualitatively assess the sugar pucker. Exchangeable protons were assigned from $^{1}H,^{1}H$-NOESY spectra recorded in H_2O/D_2O (90:10) at different temperatures (275, 280, 282, 285 K) with 150 ms mixing times, employing a 1-1 spin-echo pulse sequence or the excitation sculpting pulse scheme for water suppression.

Heteronuclear 2D experiments were performed on $^{13}C,^{15}N$ labelled RNA samples. $^{1}H,^{13}C$-HSQC experiments were employed to attribute sugar and aromatic carbons. Two separate experiments were usually recorded for the aromatic and the aliphatic range of the ^{13}C resonances. $^{1}H,^{15}N$-HSQC experiments were used to attribute the imino nitrogens. SOFAST sequences were employed for these experiments[409]. ^2J ^{1}H,^{15}N-HSQC spectra were recorded to assign purine N7 and adenine N1 and N3[391,410]. 3D ^{1}H,^{1}H-NOESY-^{1}H,^{13}C-HSQC were also performed on D1kz[410].

^{31}P chemical shifts and correlations to sugar puckers were determined in ^{31}P 1D and ^{1}H,^{31}P-TOCSY spectra[411].

The base-pairing pattern was established from J_{NN} HNN-COSY spectra, correlating the imino nitrogens of uracil and guanine across the hydrogen bond to the N1 of adenine or the N3 of cytosine on the other side of the double helix[58,389].

Phase sensitive 2D w_1,w_2-^{13}C,^{15}N-filtered NOESY and TOCSY experiments with watergate H_2O suppression using nucleotide-specifically ^{13}C,^{15}N-labelled samples helped to reduce spectral overlap[412]. An X-double-half-filtered experiment[56] was used to confirm dimerization of D1kz on a sample containing both unlabelled and ^{13}C,^{15}N labelled D1kz. This experiment yields four subspectra (w_1,w_2-filtered, w_1,w_2-edited, w_1-filtered,w_2-edited, w_1-edited,w_2-filtered), allowing to distinguish intramolecular from intermolecular cross peaks.

NMR data were processed using TopSpin 3.0 (Bruker) and analysed using Sparky (http://www.cgl.ucsf.edu/home/sparky/). The 3D experiments were analysed using CARA (http://cara.nmr.ch/doku.php). NOE peak volumes for the structure calculation were integrated with the Gaussian peak-fitting function in Sparky.

RDC measurements. Partial alignment for measuring residual dipolar couplings (RDCs) was achieved by adding Pf1 filamentous bacteriophage (*D1kz:* ~8.0 mg/mL, *AvD5:* 25 mg/mL) to ^{13}C,^{15}N labelled RNA samples. In the case of D1kz, where RDCs were collected in the presence of $MgCl_2$, the phages were spun down for 2 h at 50000 rpm at 4°C prior to usage to exchange the phosphate buffer with a 60 mM KCl water solution, to avoid M^{2+}-phosphate complexation. D1kz RDCs were measured employing J-modulated ^{1}H,^{13}C-HSQC experiments[413] in the presence of 7 mM $MgCl_2$, run separately on aromatic and sugar regions. NOE peak volumes were integrated using Topspin 3.0 (Bruker) or CCPN (www.ccpn.ac.uk/ccpn). The data were fitted using either Origin® or gnuplot (http://www.gnuplot.info/). AvD5 RDCs were determined from the doublet peak separation in ^{1}H,^{13}C- and ^{1}H,^{15}N-HSQC spectra recorded without heteronuclear decoupling on isotropic and partially aligned samples.

4.4 Interaction studies of D1kz with D5 from Sc.ai5γ

To investigate the interaction of D1kz and D5 from Sc.ai5γ in solution we tried several different approaches, none of which yielded any sign of an interaction between the two constructs. We varied the ratio of D1kz and D5 as well as Mg(II) concentration and used differently labelled samples. First D1kz and a completely deuterated D5 were mixed 1:1 (0.5 mM each) and ^{1}H,^{1}H-NOESY spectra were measured in the absence and presence of 1.5 mM and 4.8 mM Mg(II) at 300 K to monitor changes in D1kz that would indicate interaction with D5. In a second attempt D1kz (0.6 mM, unlabelled) was titrated with 0.5, 1 and 2 equivalents of D5 (^{13}C,^{15}N labelled) in the presence of 10 equivalents of Mg(II) at 300 K. Changes in D5 were followed by ^{1}H,^{13}C- and ^{1}H,^{15}N-HSQC spectra, while changes in D1kz would have been observable in recorded w_1,w_2-^{13}C,^{15}N-filtered NOESY spectra. In a further experiment we used a labelled D1kz sample (0.3 mM) and added up to 10 equivalents of D5, followed by addition of 2.5, 3, 4, 8, 16 and 20 mM Mg(II). Prior to recording ^{1}H,^{13}C- and ^{1}H,^{15}N-HSQC spectra at 290 K the sample was heated to 80°C to separate possible D1kz dimers. Finally we mixed ^{13}C,^{15}N labelled D1kz (0.5 mM) with unlabelled D5 (0.35 mM) in the presence of 12 mM Mg(II) and a ^{1}H,^{13}C-HSQC spectrum as well as an X-double-half-filtered NOESY experiment[56] were recorded at 300 K.

4.5 Structure calculations

Restraint collection. Distance restraints were derived from ^{1}H,^{1}H-NOESY spectra measured in D_2O at 60 and 250 ms mixing time and from ^{1}H,^{1}H-NOESY spectra measured in H_2O/D_2O (9:1) at 150 ms mixing time. Not or only moderately overlapping peaks were integrated in Sparky (http://www.cgl.ucsf.edu/home/sparky/) and distances were calibrated to pyrimidine H5H6 and H1'H2' peak volumes using DYANA's CALIBA macro[414]. Based on these, all assignable peaks were categorized as strong (1.8-3 Å), medium (1.8-4.5 Å), weak (3-6 Å) and very weak (4-7 Å). In addition, dihedral angle restraints were employed: For bases in stable base pairs and helical regions χ was restrained to *anti*-conformation ($-160° \pm 20°$). In D1kz, G32 was restrained to *syn*-conformation

(60° ± 20°) owing to its strong intranucleotide H8-H1' peak and U29, G15, A16, A17 and A18 were restrained to *anti* according to literature data[281,387]. For the same reasons G16 in D1-27 was restrained to *syn*. For all other bases in D1kz, AvD5 and D1-27, χ was left unrestrained. Sugar puckers were restrained to C3'-endo conformation ($\delta = 85°$, $\nu_1 = -25°$, $\nu_2 = 37.3°$, ± 10°) in helical regions, to C2'-endo conformation for residues showing strong H1'H2' and medium H1'H3' cross peaks in 50 ms ^{1}H,^{1}H-TOCSY spectra ($\delta = 145°$, $\nu_1 = 25°$, $\nu_2 = -35°$, ± 10°) (A6, A16, U30, C31 and A43 in D1kz; G17, U18 in AvD5; A6, U13, C14, A21 in D1-27)[281,385,386,415] and left unrestrained in all other cases. Backbone angles were restrained to standard A-RNA values ($\alpha = -68°$, $\beta = 178°$, $\gamma = 54°$, $\epsilon = -153°$, $\zeta = -71°$, ± 20°) in regions where the NOESY peak pattern and base-pairing strongly suggest standard helical conformation.

In AvD5 some further dihedral angles could be restrained weakly: α and ζ were restrained to 0 ± 120° in the non-helical regions based on the fact that all ^{31}P signals are in the range of −1.2 to 0 ppm with respect to external H_3PO_4 (85%), which indicates that no α or ζ angles are in the trans range[51]. β was restrained weakly to the trans region (120-240°) when the P-H5' and P-H5" peaks were clearly weak or absent in ^{1}H,^{31}P-TOCSY spectra (A16, G17, U18, U19). In the non-helical regions, γ angles were judged qualitatively from the H4'-H5" and H4'-H5' peaks in the 50 ms TOCSY. If both H5'/H5" to H4' couplings are small, γ is gauche$^+$ or around 240° (U15, A16, U18 were restrained to −130-90°). γ is gauche$^-$ or trans when one of the two is strong (U19, restrained to 120-330°)[416–418]. G17-γ was left unrestrained because both H5' and H5" to H4' peaks are visible but not very strong.

All other backbone dihedral angles in D1kz, D1-27 and AvD5 were left unrestrained.

Hydrogen bonds in W-C base pairs were kept stable by short distance restraints between the hydrogen and the acceptor as well as between the donor and the acceptor and by planarity restraints.

Calculation protocol. Starting from an extended conformation an initial ensemble of 100 structures was calculated using CNS 1.21[419,420] (40 ps torsion angle dynamics at 20000 K, followed by slow cooling for 90 ps in torsion angle space and 30 ps in cartesian space). This was followed by 60 to 80 ps of refinement using Xplor-NIH 2.24[348,349], slowly cooling from 3000 to 50 K, during which the single-bond ^{1}H,^{13}C RDC restraints were introduced (with a force constant increasing from 0.01 to 1.0 kcal mol^{-1} Hz^{-2}). Axial component and rhombicity of the alignment tensor were determined as −20.8/0.2 for D1kz and −26.9/0.2 for AvD5 in a gridsearch[421] using Xplor-NIH[348,349]. RDC error limits were set to generous 1.5 Hz, to allow for the uncertainty introduced by the fact that peaks are broadened in the presence the Pf1 phages.

Additional sets of structures were refined in Xplor-NIH[348,349] including weak distance restraints (6.0 Å to the central Co(III) ion) to two and three cobalt(III)-hexammine molecules, respectively, in AvD5 and D1kz.

Structure analysis The twenty lowest-energy structures (none of them had NOE violations >0.2 Å , dihedral angle violations >5° or RDC violations >3 Hz) were selected as the final ensemble and analysed using MolMol[422] and molprobity (http://molprobity.biochem.duke.edu/). Electrostatic surface potentials were determined using APBS (http://www.poissonboltzmann.org/apbs)[423] and visualized in Pymol (http://www.pymol.org).

4.6 UV melting studies

AvD5 samples (1 µM in 60 mM KCl, 10 µM EDTA) were thermally denatured in 3-4 consecutive cycles of heating and cooling from 20°C to 90°C and back at a rate of 0.5-1°C/min. Prior to measurement samples were carefully covered with paraffin oil and the cuvettes were sealed with parafilm to avoid evaporation. Absorption values at 260 nm were recorded in steps of 0.5°C. The melting curves and their slopes (as determined in spreadsheets) were plotted using gnuplot (http://www.gnuplot.info/). Melting temperatures were assessed using Levenberg-Marquardt nonlinear least-squares regression and an equation derived from van't Hoff analysis[424,425], describing one melting transition with linearly increasing baselines in the folded and unfolded states:

$$f(T) = \frac{bf + mf * T + (bu + mu * T) * e^a}{1 + e^a} \tag{4.1}$$

where

$$a = \Delta H * \frac{\frac{1}{T_m} - \frac{1}{T}}{8.31441} \tag{4.2}$$

and T_m is the melting temperature in K, ΔH is the enthalpy of unfolding and bf/mf and bu/mu are the slopes/intercepts of the baselines in the folded and unfolded states, respectively.

4.7 Fluorescence spectroscopy

All experiments were performed on 1 µM AvD5-AP24 samples. The excitation wavelength was set to 305 nm. Emission scans were performed from 330 to 520 nm with a scan speed of 100 nm min^{-1}. Excitation and emission slits were set to 10 and 20 nm, respectively. The pH dependence of AP24 fluorescence intensity was determined by adding KOH in 14 steps going from pH 5 to pH 9, followed by adding HCl in 16 steps returning to a pH of 5. At each step the pH was left to stabilize for 10-20 min. For details on the Mg(II) titrations of AvD5-AP24 which were also followed by fluorescence, see the following section.

4.8 Metal titrations

To determine K_D values the amount of added metal ion was plotted against changes in chemical shift or fluorescence. Clearly monophasic curves were fit to a single binding isotherm (Equation 4.3)[114,140] by Levenberg-Marquardt non-linear least-squares regression using gnuplot (http://www.gnuplot.info/):

$$\begin{aligned} f(x) = {} & dF_1 + (dB_1 - dF_1) * \\ & * \frac{([RNA] + x + K_D) - \sqrt{([RNA] + x + K_D)^2 - 4 * [RNA] * x}}{2 * [RNA]} \end{aligned} \tag{4.3}$$

where dB_1 and dF_1 are the measured quantity — chemical shift in NMR experiments and intensity in fluorescence experiments — of the unbound and fully bound species, respectively, and [RNA] is the concentration of the RNA in mM. For biphasic curves the equation was extended by an additional term for a second independent binding event :

$$
\begin{aligned}
f(x) = dF_1 + (dB_1 - dF_1) * \\
* \frac{([RNA] + x + K_{D,1}) - \sqrt{([RNA] + x + K_{D,1})^2 - 4 * [RNA] * x}}{2 * [RNA]} - \\
- dB_2 * \frac{([RNA] + x + K_{D,2}) - \sqrt{([RNA] + x + K_{D,2})^2 - 4 * [RNA] * x}}{2 * [RNA]}
\end{aligned}
\tag{4.4}
$$

Metal ion titrations followed in NMR spectra. All metal titrations that were followed by recording NMR spectra were performed by the stepwise addition of small volumes (µLs) of the metal ion stock solution to the RNA sample (250-300 µL) in the NMR tube. After addition the sample was rigorously mixed and then spun down using a hand-centrifuge. The volume increase (up to 50 µL) was taken into account in the metal ion concentrations, but not the concomitant dilution of the RNA.

RNA samples were usually 0.5 to 1 mM in a solution of 60 mM monovalent salt and 10 µM EDTA (details are given in the legends of the respective graphs and tables in the Results part and in the Appendices). Chemical shift changes were followed in ^{1}H 1D, ^{1}H,^{1}H-NOESY, ^{1}H,^{13}C- and ^{1}H,^{15}N-HSQC spectra.

Mg(II) titrations were performed by the stepwise addition (0.5, 1 or 2 mM steps) of a 0.1 M $MgCl_2$ stock solution up to a maximum concentration of 12 mM Mg(II). Cd(II) titrations were performed in steps of 0.3-0.5 mM up to a maximum concentration of 3-4 mM Cd(II), using a 0.1 M $Cd(ClO_4)_2$ stock solution. Only in the case of Cd(II) $KClO_4$ was used instead of KCl as the monovalent salt in the sample because Cl^- tends to associate with Cd(II) to form $CdCl^+$ (B. Knobloch and R. K. O. Sigel, unpublished results). A 0.1 M stock solution of $Co(NH_3)_6Cl_3$ was titrated to NMR samples in steps of 0.4-0.5 mM up to final concentrations of 3-4 mM. The Mn(II) titration on AvD5 was performed with a 0.01 M $MnCl_2$ stock solution, which was added in steps of 15 to 30 µM, up to a final concentration of 120 µM Mn(II). KCl titrations were performed using a 2 M stock solution in steps of 50-100 mM up to a maximum concentration of 400 mM.

Mg(II) titrations followed by 2-aminopurine fluorescence. Fluorescence experiments were performed on 1 µM AvD5-AP24 samples in the presence of 0.001, 0.0025, 0.005, 0.0075, 0.01, 0.025, 0.05, 0.075, 0.1, 0.25, 0.5, 0.75, 1, 2.5, 5, 7.5, 10, 25, 50, 75 and 100 mM $MgCl_2$. $MgCl_2$ stock solutions were prepared at 1, 10, 100, 200 and 1000 mM to minimize the added volume. To avoid diluting the RNA, at every titration step we added not only the required volume of $MgCl_2$ stock solution but in addition the same volume of a 2x concentrated AvD5-AP24 solution. Prior to measurement the solution was mixed by gently pipetting up and down 3-5 times in the cuvette. Experiments were performed in the absence and in the presence of 10 or 40 mM MOPS buffer. The pH of samples and stock solutions was adjusted using a minimum of HCl and KOH. The thereby added K(I) was taken into account in the total K(I) concentration of the samples.

4.9 Determination of pK_a values

Two series of experiments, one using ^{1}H,^{1}H-NOESY spectra on an unlabelled AvD5 sample and one using ^{1}H,^{13}C-HSQC spectra on a ^{13}C,^{15}N labelled AvD5 sample, were recorded at pHs between 5.5 and 9. Least-squares fits to a Henderson-Hasselbalch-derived equation[426] (Equation 4.5) were performed for the chemical shift changes as a function of pH of all unoverlapped H1', H2, H5, H6, H8, C2, C6 and C8 resonances in the bulge and catalytic triad, respectively, that do not disappear too fast to allow a reliable fit.

$$\delta(pH) = \frac{10^{pK_a-pH} * \delta AH + \delta A}{1 + 10^{pK_a-pH}} \tag{4.5}$$

where δ(pH) is the observed chemical shift in ppm at a certain pH, δAH is the chemical shift of the completely protonated species and δA is the chemical shift of the unprotonated species.

5

Summary

Group II introns[197] are large modular ribozymes. In a process known as splicing they catalyze their own excision from RNA strands and the joining of the flanking exon sequences. Of their six domains, two stand out in particular: domain 1 (D1), which is the largest one, providing a scaffold to which the other domains dock, and the short hairpin domain 5 (D5) where most of the catalytically essential functionalities are concentrated[75,263,268–270,272]. Those two domains alone retain some catalytic proficiency for the cleavage reaction at the 5'-splice site[254].

Group II intron splicing is accomplished in two consecutive phosphotransesterification steps, both of which are dependent on divalent metal ions. Solution studies have long indicated a catalytically important Mg(II) binding site close to a two-nucleotide bulge and an adjacent conserved AGC triad in D5[75,256,281,282] and in a recent crystal structure of a group IIC intron two Mg(II) ions are found coordinating there[106,236]. The same catalytic mechanism is used by the eukaryotic spliceosome in nuclear pre-mRNA processing[162,165]. The spliceosome is a megadalton machinery involving a multitude of proteins, but whose catalytic activity is located in an RNA component with striking similarities to D5: in the internal stemloop (U6 ISL) of the U2-U6 snRNA complex[163].

Divalent metal ions are not only essential for catalysis, they are also required for the folding of group II introns to their native, catalytically active structure[244,307]. The first step of folding is the compaction of D1, which has been shown to depend heavily on functionalities situated in the small κ-ζ element of D1[258,320]. The same small element co-incidentally also provides the docking site for D5 in the later stages of folding.

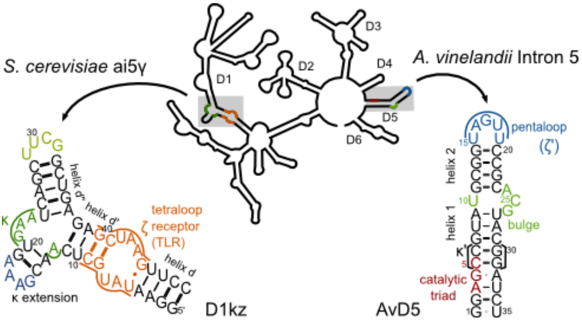


Figure 5.1 Secondary structure schemes of AvD5 and D1kz, indicating their position in a group II intron secondary structure and important structural elements. Greek letters indicate tertiary interaction sites between D1 and D5.

This work comprises two self-contained studies focusing on the investigation of D5 and the κ-ζ region of D1, the two regions that are at the heart of group II intron catalysis and folding, respectively, and their interplay with essential metal ion co-factors.

In the first part, the solution structure of D5 from *A. vinelandii* Intron 5 (AvD5, from intron Av.I5) was solved. This is the first structural study on a group II intron of the bacterial class E. Two interesting features distinguish D5 in this intron class from previously studied D5s. Not only is the more common capping GNRA tetraloop replaced by a U-rich pentaloop which has never been studied so far, but also the G-U pair at the most conserved position in the "catalytic triad" region is replaced by a G-A mismatch (G4-A32 in Figure 5.1). Both features play an important role in tertiary contacts of D5 to other parts of the intron and the G-A pair is furthermore involved in the active site metal ion coordination. To determine the impact of these two characteristic variations observed in bacterial class E group II introns, solution NMR structural studies were complemented by detailed metal ion binding studies.

The second part of this work is concerned with the NMR investigation of a 49-nt construct (D1kz, Figure 5.1) containing the κ-ζ region of *S. cerevisiae* group II intron ai5γ (Sc.ai5γ). By studying it in solution and in the absence and presence of its natural co-factor Mg(II) we were aiming to elucidate the mechanistic basis of the first rate-limiting step of group II intron folding, which has been assigned to this small region in previous studies[258,320]. We solved the structure by NMR spectroscopy and investigated the reason for the strong dependence on divalent metal ions of the first step of group II intron folding.

AvD5 – Strongly perturbed pK_as, solution structure and metal ion binding

AvD5 is a 35-nt hairpin with three unpaired regions: a G-A mismatch in the catalytic triad, an asymmetric internal loop referred to as the "bulge" and a terminal pentaloop (Figure 5.1). All three of them proved challenging in their own way in the process of structure determination and revealed interesting features that further our understanding of the assembly of D5 into the folded intron.

The most conserved position in D5 of group II introns is the G at the center of the catalytic triad (G4 in Figure 5.2), which is almost exclusively paired with a U. Only some members of the bacterial intron class E[212,234], among which is *A. vinelandii* Intron 5, have a G-A mismatch at this position. All previously solved D5 structures[106,236–238,280–282] have a G-U wobble at this position and the catalytic triad forms a regular base-paired helix. By NMR spectroscopy we could show that the pK_a of the A-N1 in the G-A mismatch is significantly perturbed to a value of 6.69 ± 0.10 (in 60 mM KCl). Nucleobases with strongly shifted pK_as have been discovered in the active sites of many ribozymes[14,181,427]. Protonation in the neutral pH range can be directly involved in acid-base catalysis, but can also be important to maintain conformational flexibility. In the case of AvD5, we show that the protonation of the A-N1 position induces a switch of the opposing G from an *anti-* to a *syn*-conformation (Figure 5.2). This exposes the Watson-Crick face of the G and would make it accessible for tertiary interaction partners. The conformational change, but also the increased mobility evoked by the protonation equilibrium at neutral pH could be implicated in the first or second step of splicing in AvD5. It might also play a role in the rearrangement of the active site between the two splicing steps required for the exchange of substrates.

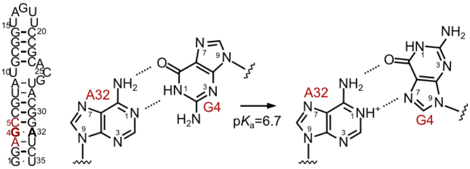


Figure 5.2 Protonation at A32N1 leads to a *anti-syn* switch of the opposing G4 in AvD5. G4 and A32 are marked in bold in the secondary structure on the right, the catalytic triad is marked in red.

A second particularity of D5 in class E introns is the terminal pentaloop[212,234].

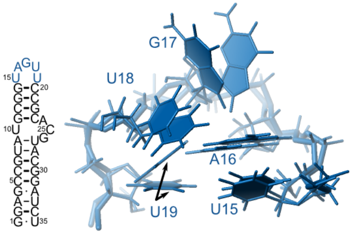


Figure 5.3 Conformation of the UAGUU pentaloop of AvD5 in the two lowest energy structures of the structure ensemble. Loop residues shown on the right are marked in blue in the secondary structure scheme. Black arrows mark the two distinct conformations of U19, bonding with either U15 or A16, respectively.

The majority of group II introns contains a tetraloop at this position — most often one of the common GNRA-type loops — which docks into a conserved receptor in D1. In AvD5 the so far undescribed U-rich pentaloop turned out to be a more variable structure. Two nucleotides at the top (G17 and U18) are mostly exposed to the solvent and strong intranucleotide base-to-sugar NOEs incompatible with one single structure strongly indicate fast conformational exchange. U15 and A16 stack onto helix 2 and U19 alternately pairs with either of them in the final NMR structure ensemble (Figure 5.3). In addition to the indicated dynamics, also the similarities to the loop of the IRE (iron responsive element) hairpin, which is known to completely rearrange upon docking to its interaction partner[378,384], suggest that upon binding to its receptor the AvD5 loop will undergo a conformational change.

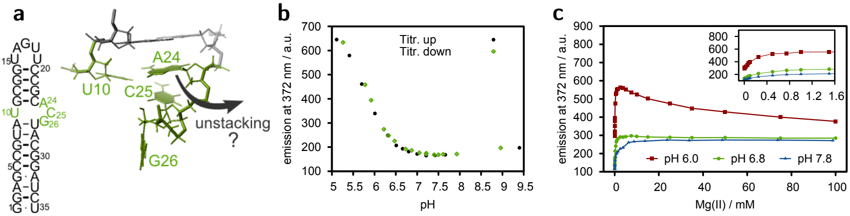


Figure 5.4 pH and Mg(II) dependence in the AvD5 bulge. a) The conformation of the bulge in AvD5 at pH>7 with a stacked-in A24. Bulge nucleotides are marked in green. b) Fluorescence emission of a 2-aminopurine (AP) in the place of A24 increases only slightly between pH 7 and pH 6.5, strong increase starts only below pH 6. The curves correspond to two separate titrations, one starting at high pH, one starting at low pH. c) AP fluorescence increase at three different pHs upon addition of Mg(II) indicating unstacking motions of the base. The increase is pH dependent. The small inset graph shows a zoom of the first data points of the larger graph.

The bulge, the third non-canonical region in D5 apart from the catalytic triad and the loop, is doubtlessly one of the most well-studied elements in group II introns. Its structure has been solved several times in different constructs demonstrating that crystal and solution conformations are not the same[106,236–238,280–282]. The AvD5 structure is no exception, showing A24, and mostly also C25, stacked inside and G26 solvent-exposed (Figure 5.4a), which is completely opposite of what is observed in crystal structures, where G26 pairs with U10 and A24 and C25 are exposed. However, in our NMR spectra we noticed a distinct pH dependence that was not described before. Resonances broaden out and completely disappear below pH 6 and a pK_a of 6.38 ± 0.12 close to neutral pH was determined in the bulge. Surprisingly, the chemical shift changes of N1 and C2 resonances indicate that not the nucleotides in the bulge, but the adjacent A9-N1 is being protonated. While crystal structures show A9 paired with U27, which would stabilize the unprotonated form of A9, the base pair appears to be unstable in solution and there are interactions with bulge nucleotides which could facilitate protonation.

Broadening and disappearance of bulge resonances in NMR spectra indicates a conformational exchange. To determine if this involves the unstacking of the A24 base — which would be one step towards the conformation observed in the fully assembled group II intron — we performed bulk fluorescence experiments on a construct containing the fluorescent nucleobase analogue 2-aminopurine (AP) at bulge position 24 (AvD5-AP24). The results indicate an increase in solvent exposure of this nucleotide but not with the pK_a value of 6.4 of A9 but rather with a pK_a<6 (Figure 5.4b). The pH dependence is reminiscent of spliceosomal U6 ISL, where the bulge nucleotide at the equivalent position turns from a stacking to a solvent exposed position at low pH[298]. In addition, fluorescence measurements indicate a Mg(II) ion binding with high affinity (K_D = 20-800 µM, depending on the salt concentration) at the AvD5 bulge, which also leads to

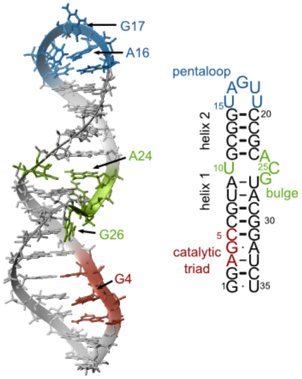


Figure 5.5 Representative solution structure at pH 5.2 and secondary structure scheme of AvD5. Nucleotides mentioned in the text are indicated in the structure. Pentaloop: blue, bulge: green, catalytic triad: red.

increased solvent exposure of AP24 (Figure 5.4c), biasing the bulge towards the intron-docked conformation. The affinity is pH dependent, with Mg(II) binding about 4x more strongly at pH 6 than at pH 7.8.

We have solved the solution structure of D5 (Figure 5.5) of *A. vinelandii* Intron 5 under two different pH conditions. NOE-derived distance restraints for the catalytic triad region were collected at pH 5.2 and 7.8. For the rest of the hairpin only high-pH restraints were used, because there is either no significant change with pH (pentaloop) or the peaks are not observable at low pH (bulge). In this way we derived two structure ensembles with RMSDs of 2.65 ± 0.62 Å and 3.43 ± 0.91 Å for pH 5.2 and 7.8, respectively, which differ only at the catalytic triad. At low pH a G_{syn}-AH^{+}_{anti} base pair is formed, whereas the region is very disordered at high pH with G4 in different conformations inside the helix or solvent-exposed. Even though a small number of residual dipolar couplings (RDCs) were employed, the bulge still acts as a hinge around which the helices can bend to some degree, causing the comparatively high overall RMSDs.

NMR experiments were employed not only for structure determination but also to identify metal ion binding sites in AvD5 that could play both structural and catalytic roles. The high affinity Mg(II) binding site at the bulge determined in fluorescence studies is apparent only by an initial change of direction of the chemical shift perturbations (Figure 5.6). ^{13}C and ^{15}N chemical shift changes do not indicate any strong direct binding to the nucleobases, rather suggesting that this Mg(II) coordinates to phosphate oxygens of the backbone in the bulge. In addition to this stronger binding site, we identify 4-5 weaker metal ion binding sites from the Mg(II) dependent ^{1}H chemical shifts, including a second one in the region of the bulge. All determined affinities are in the low to medium

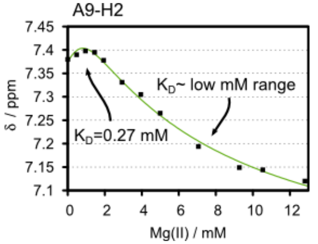


Figure 5.6 Chemical shift perturbation of a resonance close to the bulge in response to Mg(II) as observed in ^{1}H,^{1}H-NOESY spectra of AvD5 at pD 6.8. The graph illustrates how the observed biphasic behaviour can be described by two binding sites in the submillimolar and low millimolar range. The number of data points is not sufficient for a more quantitative analysis of the two K_Ds.

millimolar range, but the lowest affinities are found in the pentaloop. Cd(II) and cobalt(III)hexammine were employed as mimics for Mg(II) for inner- and outer sphere coordination sites, respectively. Both ^{13}C and ^{15}N chemical shift perturbations in the presence of Cd(II) and intermolecular NOEs from the RNA to cobalt(III)hexammine show a preference for binding at the tandem G-C pairs in helix 1 and helix 2. The pentaloop does not bind any of the tested metal ions strongly.

D1kz – The solution structure and its dependence on di- or trivalent metal ions

Our 49-nt D1kz construct contains the κ-ζ region of Sc.ai5γ. ζ is an 11-nt tetraloop receptor, a motif which has been studied previously in solution[385,386], revealing a conformational change between its free and tetraloop-bound states. In D1kz we find the ζ region sensitive to metal ions, manifested in a broadening and disappearance of many peaks in increasing metal ion concentrations. In addition, we observe a transient intermolecular interaction of the GAAA-tetraloop of one D1kz molecule with ζ from another molecule above ~4 mM Mg(II), indicating formation of a homodimer. The receptor, however, remains in its "unbound" conformation as indicated by a lack of characteristic spectral features, suggesting that the conserved canonical interaction of the 11-nt-receptor with the GAAA tetraloop in the transient homodimer is prevented by sterical hindrances.

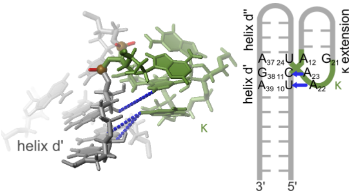


Figure 5.7 The interaction of κ (green) with helix d' in D1kz. Some observed NOEs are indicated in blue. Nucleotides that are shown on the left are labelled in the scheme on the right where the interaction is indicated in the same colours.

κ is a 4-nt motif embedded in a three-way junction, which has been shown to form a GNRA-like structure when bound to D5 in the complete intron[237,266]. Our NMR experiments indicate that the GNRA-fold is adopted already in the absence of D5 and divalent metal ions. In addition, it forms A-minor interactions with the adjacent helix d' (Figure 5.7). This feature is remarkable because it is not observed in the folded intron, where κ is binding to κ' in D5[236,266] (Figure 5.8). Addition of ScD5 (D5 from Sc.ai5γ) to this D1kz construct in solution is not sufficient to release the intramolecular A-minor interactions. Even when

ScD5 is added in excess and in the presence of Mg(II), we do not observe any sign of interaction with D1kz. We therefore suspect that the network of tertiary interactions in the frame of the folded intron, possibly involving the adjacent coordination loop (Figure 5.8b), might be required to mold the κ-ζ region and D5 into their binding-competent forms.

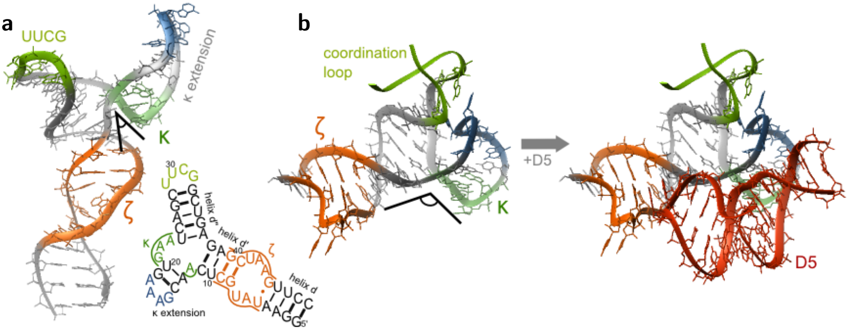


Figure 5.8 The different orientation of the κ extension in solution (the D1kz structure solved in this work) (a) and in the κ-ζ region of a group IIC crystal structure (b) (PDB-ID: 3G78[237]). In (a) A-minor interactions connect κ to helix d' making the angle between the two small, while these interactions are disrupted in (b), allowing D5 to dock. The latter conformation might be facilitated by an interaction of the κ extension with the adjacent coordination loop.

While Mg(II) does not play a role in the formation of the observed A-minor interactions, it proved to be an essential stabilizing factor for the base pairs adjacent to the κ three-way junction. In the presence of only monovalent salt, the neighbouring base pairs are not stable, which is indicated by the absence of observable imino protons. In addition, there are gaps in the sequential walk in the region where the strands entering and leaving the κ extension cross (C11, A12, A37, G38). Upon addition of Mg(II) the base pairs are stabilized and the missing peaks appear. Also a cross peak between U24H3 and G38H1 emerges, connecting helix d' and d" across the three-way junction. This strongly indicates that co-axial stacking is established between the two helices, a feature which is also observed in crystal structures[236,280]. The Mg(II)-dependent stabilization of the κ three-way junction thus biases the structure toward the conformation observed in the folded intron, evocative of the metal dependence of the initial compaction of D1, which also involves the stabilization of helices in the region of κ and ζ[258,320].

Employing distance restraints collected in ^{1}H,^{1}H-NOESY spectra in the presence of Mg(II), we derived a structure ensemble with an RMSD of 4.30 ± 1.23 Å in restrained molecular dynamics simulations (Figure 5.8a). Visualization of the surface potential at the three-way junction reveals a region of negative potential at the phosphates of A12 and U24 (Figure 5.9), indicating a favourable binding pocket for metal ions. The stabilization in the κ region is accompanied by strong perturbations of the ^{1}H resonances. It is, however, impossible to distinguish between the direct effects of metal ion binding and the effects caused by structural rearrangement, making it difficult to localize the binding site precisely. In the chemical shift perturbations of ^{13}C and ^{15}N there is no evidence for a direct coordination to the nucleobases, which supports the idea that Mg(II) coordinates at the phosphate backbone.

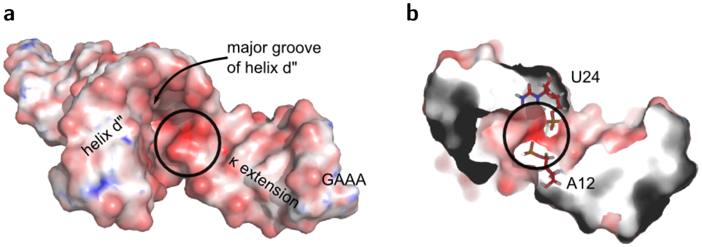


Figure 5.9 a) A patch of high negative electrostatic surface potential between helix d” and the κ extension denotes the location of A12 and U24 phosphates (b). The potential is represented by a colour gradient from red (-1068 mV) to blue (134 mV).

The observed stabilization of the κ region is not specific for Mg(II), in fact Cd(II) and cobalt(III)hexammine are even more efficient at the task, i.e. in their case lower metal ion concentrations are sufficient. For both ions we observe signs of interaction at the major groove of helix d” via ^{13}C and ^{15}N chemical shift perturbations and intermolecular NOEs, respectively. Since it is highly unlikely that both Cd(II) and cobalt(III)hexammine would coordinate in the same manner, we attribute the stabilizing influence to the charge neutralization in a region of concentrated electronegative potential between helix d” and the κ extension helix which does not require a specific coordination mode (Figure 5.9).

Conclusion

The two RNA constructs investigated in this thesis are only small parts of the large group II intron ribozymes from which they are taken. In both cases it is apparent that the structures in solution differ in many ways from the ones in the crystal structure and in the folded intron. In particular the non-canonical, unpaired regions show signs of different degrees of flexibility. Nonetheless, the natural co-factor Mg(II), which is essential for group II intron folding and catalysis, is attracted to the same binding sites in solution as in the fully assembled intron. Mg(II) induces changes in the conformation and flexibility of specific substructures, biasing both the AvD5 bulge as well as the region around the D1kz three-way junction towards the conformation observed in the folded group II intron. Our data thus indicates that the assembly of large RNAs is based both on the induced fit enabled by the flexibility of non-canonical, unpaired regions and on the preselection of certain conformations by Mg(II). In this way what we observe in these two small RNA constructs can help us to understand basic principles in the folding of large RNAs.

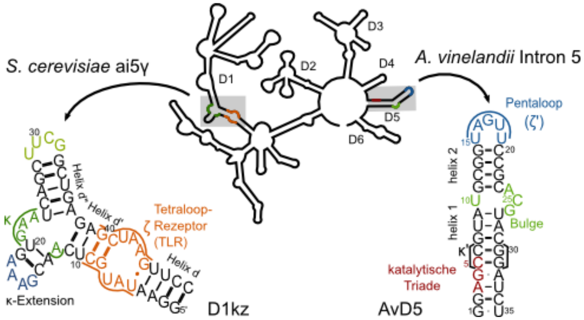
6

Zusammenfassung

Gruppe II Introns[197] sind grosse, modulare Ribonukleinsäuren (RNS). In einem als "Spleissen" bekannten Prozess schneiden sie sich aus RNS-Strängen heraus und verbinden die freiwerdenden Enden der benachbarten Exons miteinander. Zwei ihrer insgesamt sechs Domänen stechen besonders hervor: Domäne 1 (D1), die grösste, die den übrigen als Gerüst dient und die kleine Domäne 5 (D5), in welcher sich die meisten für die Katalyse essentiellen funktionellen Gruppen konzentrieren[75,263,268–270,272]. Diese beiden Domänen gemeinsam reichen aus für einen Rest katalytischer Aktivität für die Teilungsreaktion an der 5'-Spleissstelle[254].

Das Spleissen durch Gruppe II Introns erfolgt in zwei aufeinanderfolgenden Phosphotransesterifizierungsschritten, welche beide von zweiwertigen Metallionen abhängig sind. Hinweise auf eine Mg(II)-Bindungsstelle in der Nähe der ungepaarten Nukleotide die den "Bulge" von D5 bilden und der sogenannten katalytischen Triade (Figure 6.1) gab es schon lange aus Studien in Lösung[75,256,281,282]. In kürzlich gelösten Kristallstrukturen eines Gruppe IIC Introns wurden dort auch wirklich zwei koordinierte Mg(II)-Ionen gefunden[106,236]. Derselbe katalytische Mechanismus wird auch vom eukaryotischen Spleissosom bei der Prozessierung der pre-mRNA verwendet[162,165]. Das Spleissosom ist eine Megadalton-Maschinerie, die auch eine Vielzahl von Proteinen mit einschliesst, deren katalytische Aktivität aber in einer RNS-Komponente mit auffallender Ähnlichkeit zu D5 steckt: im "internal stemloop" (U6 ISL) des U2-U6 snRNA Komplexes[163].

Zweiwertige Metallionen sind nicht nur für die Katalyse essentiell, sie werden auch für die Faltung der Gruppe II Introns in ihre native, katalytisch aktive Struktur benötigt[244,307]. Der erste Schritt bei der Faltung ist eine Kompaktierung von D1, welche besonders von Funktionalitäten im kleinen κ-ζ-Element von D1 bestimmt wird[258,320] (Figure 6.1). Dasselbe kleine Element bildet auch die Andockstelle für D5 in späteren Stadien der Faltung.



Figur 6.1 Sekundärstrukturen von AvD5 und D1kz. Ihre Position in einem Gruppe II Intron und wichtige Strukturelemente sind eingezeichnet. Griechische Buchstaben bezeichnen Tertiärkontakte zwischen D1 und D5.

Die vorliegenden Arbeit besteht aus zwei eigenständigen Studien, die sich mit D5 und der κ-ζ-Region von D1 befassen, zwei Regionen die im Zentrum der Gruppe II-Intron-Katalyse und -Faltung stehen. Wir untersuchen ihre Struktur sowie auch ihr Zusammenspiel mit essentiellen Metall-Kofaktoren.

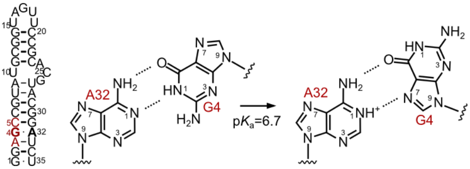
Im ersten Teil wurde die Struktur von D5 des *A. vinelandii* Introns 5 (AvD5, vom Intron Av.I5) in Lösung aufgeklärt. Es ist die erste strukturelle Studie an einem Gruppe II Intron der bakteriellen Klasse E. D5 dieser Klasse unterscheiden sich von den bisher untersuchten D5s in zweifacher Hinsicht. Einerseits wird die Haarnadelstruktur nicht wie gewöhnlicherweise von einem Tetraloop des weit verbreiteten GNRA-Typs, sondern von einem bisher nicht untersuchten Uracil-reichen Pentaloop abgeschlossen. Andererseits ist das G-U Wobble-Basenpaar an der am stärksten konservierten Stelle der katalytischen Triade durch einen G-A Mismatch ersetzt (G4-A32 in Figure 6.1). Sowohl der Loop als auch die katalytische Triade spielen eine wichtige Rolle für die tertiären Kontakte D5s zu anderen Teilen des Introns und das G-A Paar ist zudem an der Metallionen-Koordination im aktiven Zentrum des Ribozyms beteiligt. Um den Einfluss dieser beiden charakteristischen Variationen in Klasse E Gruppe II Introns zu untersuchen, haben wir strukturelle Studien via NMR Spektroskopie in Lösung durchgeführt und mit detaillierten Metallionenbindungsstudien ergänzt.

Der zweite Teil dieser Arbeit befasst sich mit der NMR-spektroskopischen Untersuchung eines 49 Nukleotide langen Konstrukts (D1kz, Figure 6.1), das die κ-ζ-Region des *S. cerevisiae* Gruppe II Introns ai5γ (Sc.ai5γ) umfasst. Durch Experimente in Lösung ohne und mit natürlichem Kofaktor Mg(II) wollten wir die mechanistischen Grundlagen des ersten, limitierenden Schrittes der Gruppe II Intron Faltung ergründen, welcher in Vorgängerstudien in dieser kleinen κ-ζ-Region lokalisiert worden war[258,320]. Die Struktur konnte mittels NMR Spektroskopie gelöst werden und liefert Hinweise auf die Ursachen für die starke Metallionen-Abhängigkeit des ersten Schritts der Gruppe II Intron-Faltung.

AvD5 – stark erhöhte Säurekonstanten, die Struktur in Lösung und ihre Metallionenbindungsstellen

AvD5 ist eine 35-nt Haarnadelstruktur mit drei ungepaarten Regionen: einem G-A Mismatch in der katalytischen Triade, einem asymmetrischen internen Loop ("Bulge") und einem terminalen Pentaloop (Figure 6.1). Im Zuge der Strukturaufklärung barg jede der drei Regionen spezifische Herausforderungen und enthüllte interessante Details, die unser Verständnis davon, wie D5 in die dreidimensinale Gestalt des gefalteten Introns eingebaut wird, erweitern.

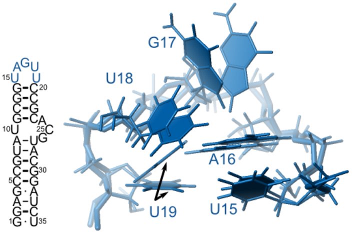
Die am stärksten konservierte Position in D5 ist das Guanin in der Mitte der katalytischen Triade (G4 in Figure 6.2), das in anderen Introns fast ausschliesslich mit einem Uracil gepaart ist. Nur Mitglieder der bakteriellen Intronklasse E[212,234], zu denen *A. vinelandii* Intron 5 gehört, haben ein G-A Mismatch an dieser Stelle. Alle bisher untersuchten D5 Strukturen[106,236–238,280–282] haben ein G-U Wobble-Paar an dieser Position und die katalytische Triade bildet eine reguläre, basengepaarte Helix. Mit Hilfe von NMR Spektroskopie konnten wir zeigen, dass der pK_a des Adenine-N1 im G-A Mismatch mit einem Wert von 6.69 ± 0.10 (in 60 mM KCl) deutlich in den neutralen pH-Bereich verschoben ist. Nukleobasen mit stark verschobenen pK_as sind schon in den katalytischen Zentren einiger Ribozyme entdeckt



Figur 6.2 Protonierung von A32N1 führt zu einem *anti-syn*-Umschalten des gegenüberliegenden G4 in AvD5. G4 und A32 sind in der Sekundärstruktur fettgedruckt, die katalytische Triade ist rot markiert.

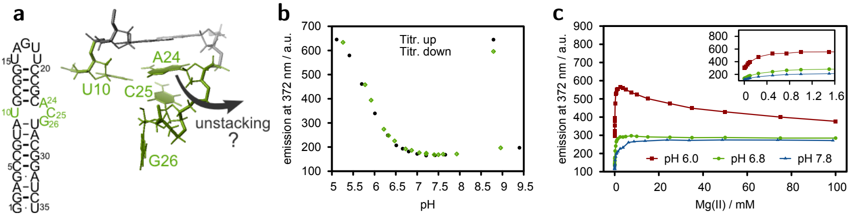
worden[14,181,427]. Protonierung im neutralen pH-Bereich kann direkt in Säure-Basen-Katalyse involviert sein, kann aber auch eine Rolle spielen, wenn es darum geht eine gewisse konformelle Flexibilität aufrecht zu erhalten. Im Fall von AvD5 können wir zeigen, dass die Protonierung von A32-N1 einen Wechsel des gegenüberliegenden G4 von einer *anti-* zu einer *syn*-Konformation zur Folge hat (Figure 6.2). Dadurch wird die Watson-Crick Seite von G4 frei und theoretisch zugänglich für tertiäre Interaktionspartner. Die Änderung der Konformation, wie auch die durch das Protonierungsgleichgewicht erhöhte Mobilität bei neutralem pH könnten bei AvD5 im ersten oder zweiten Spleissschritt beteiligt sein. Auch bei der Umlagerung im katalytischen Zentrum zwischen den beiden Schritten, die nötig ist um die Substrate auszutauschen, ist eine Rolle denkbar.

Ein zweites charakteristisches Merkmal von D5 aus Gruppe II Introns ist der terminale Pentaloop[212,234]. Die Mehrheit der Gruppe II Introns besitzt an dieser Stelle einen Tetraloop — meist einen vom weit verbreiteten GNRA-Typ — der an einen konservierten Rezeptor in D1 andockt. Im Vergleich zu jenem ist der nun zum ersten Mal untersuchte U-reiche Pentaloop in AvD5 eine flexiblere Struktur. Zwei Nukleotide am äussersten Ende des Loops (G17 und U18) sind grossteils nach außen exponiert und G17 weist mehrere miteinander unvereinbare, NOEs zwischen seiner Nukleobase und seinem Zucker auf, die auf einen konformellen Austauschprozess hinweisen. U15 und A16 sind auf Helix 2 geschichtet und U19 wechselwirkt alternierend mit jeweils einem der beiden im NMR Strukturensemble (Figure 6.3). Sowohl die angedeutete Dynamik als auch strukturelle Ähnlichkeiten zum Loop des IRE (iron responsive element) Hairpins, von welchem bekannt ist, dass er sich beim Andocken an seinen Interaktionspartner stark umordnet[378,384], lassen vermuten, dass auch der AvD5 Loop beim Binden an seinen Rezeptor seine Konformation ändern wird.



Figur 6.3 Konformation des UAGUU-Pentaloops in AvD5 in den zwei Strukturen des Strukturensembles, die die niedrigste Energie aufweisen. Loop-Nukleotide sind blau markiert in der Sekundärstruktur. Schwarze Pfeile bezeichnen die zwei unterschiedlichen Konformationen die von U19 eingenommen werden.

Der Bulge, die dritte nicht-kanonische Region in D5 nach der katalytischen



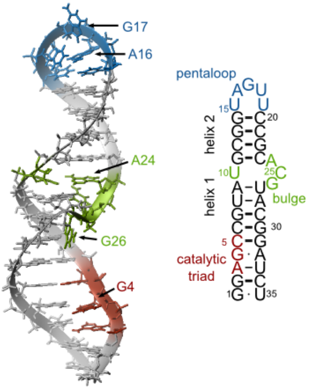
Figur 6.4 pH- und Mg(II)-Abhängigkeit im AvD5 Bulge. a) Die Konformation des Bulges in AvD5 bei pH>7 mit einem nach innen gestackten A24. Bulge Nukleotide sind grün markiert. b) Fluoreszenzemission eines 2-Aminopurin (AP) an der Position von A24 steigt nur leicht zwischen pH 7 and pH 6.5, eine stärkere Zunahme beginnt erst unterhalb von pH 6. Die Kurven wurden in zwei separaten Titrationen gemessen, eine beginnend bei hohem pH, eine beginnend bei tiefem pH. c) AP Fluoreszenzanstieg mit steigender Mg(II)-Konzentration gemessen bei drei verschiedenen pH Werten deuten. The small inset graph shows a zoom of the first data points of the larger graph.

Triade und dem Loop, ist ohne Zweifel eine der am besten studierten Regionen in Gruppe II Introns. Seine Struktur wurde mehrmals und in unterschiedlichen Konstrukten untersucht, wobei deutlich wurde, dass sich die Konformationen im Kristall und in Lösung deutlich unterscheiden[106,236–238,280–282]. Die AvD5 Struktur ist dabei keine Ausnahme mit nach innen gestackten A24 und C25 und einem nach aussen gewandten G26 (Figure 6.4a), während in den Kristallstrukturen genau umgekehrt G26 in einem Basenpaar mit U10 nach innen gerichtet ist und A24 und C25 exponiert sind. Abgesehen von der schon zuvor gut dokumentierten Struktur des Bulges, konnten wir in unseren NMR Spektren aber auch eine bislang nicht beschriebene, deutliche pH-Abhängigkeit ablesen. Resonanzen verbreitern sich und verschwinden unterhalb von pH 6 vollständig und wir konnten einen pK_a von 6.38 ± 0.12, nah am neutralen pH-Bereich, bestimmen. Die chemischen Verschiebungen von N1- und C2-Resonanzen deuten darauf hin, dass A9 — und nicht eines der ungepaarten Bulge-Nukleotide — protoniert wird. Dies ist umso überraschender als A9 in Kristallstrukturen ein Basenpaar mit U27 bildet, was zur Stabilisierung des unprotonierten A9 führen muss. In Lösung jedoch ist das Basenpaar äusserst instabil und es gibt Wechselwirkungen mit Nukleotiden aus dem Bulge die eine Protonierung begünstigen könnten.

Das Verbreitern und Verschwinden von Bulge-Resonanzen in NMR Spektren weist auf einen konformellen Austauschprozess in dieser Region hin. Um zu bestimmen, ob dieser Prozess mit einer Bewegung der A24-Nukleobase aus der

Helix in eine exponiertere Position ("Unstacking") — also in Richtung einer Konformation, die der im vollständig gefalteten Intron mehr entspricht — einhergeht haben wir Fluoreszenzmessungen durchgeführt. Mit Hilfe eines AvD5-Konstrukts (AvD5-AP24) in dem A24 durch 2-Aminopurin (AP), ein fluoreszentes Nukleobasenanalog, ersetzt war, konnten wir sehen, dass es tatsächlich zu einem Unstacking kommt. Allerdings passiert dieses nicht mit einem pK_a von 6.4, wie wir in für A9N1 bestimmt haben, sondern mit einem pK_a<6 (Figure 6.4b). Die pH-Abhängigkeit erinnert an das spleissosomale U6 ISL, wo ein Bulge-Nukleotid an der äquivalenten Position in der Haarnadelstruktur bei niedrigem pH aus der Helix ausgestossen wird[298]. In AvD5 deuten die Fluoreszenzmessungen des weiteren auf eine Mg(II)-Bindungsstelle im Bulge mit hoher Affinität hin (K_D = 20-800 µM, abhängig von der Salzkonzentration). Bindung des Mg(II)-Ions führt ebenso wie tiefer pH zu einem exponierteren AP24 (Figure 6.4c) und damit zu einer Konformation, die der im vollständigen, gefalteten Intron ähnlicher ist. Die Affinität des Mg(II) ist pH-abhängig, wobei Mg(II) bei pH 6 etwa 4x stärker bindet als bei pH 7.8.

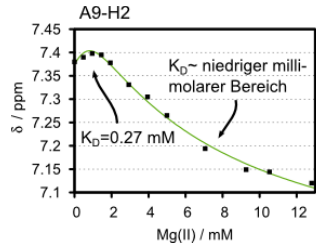
Wir haben die Struktur von D5 des *A. vinelandii* Introns 5 in Lösung bei zwei verschiedenen pHs aufgeklärt (Figure 6.5). In der katalytischen Triade wurden NOE-Intensitäten und davon abgeleitete Distanzen bei pH 5.2 und 7.8 bestimmt. Im Rest des Hairpins wurden für die Strukturberechnung nur Daten verwendet, die bei hohem pH gemessen wurden, da entweder keine signifikante pH-Abhängigkeit bestand (Pentaloop) oder die Peaks bei niedrigem pH nicht sichtbar sind (Bulge). Auf diese Weise erhielten wir zwei Strukturensembles mit RMSDs von 2.65 ± 0.62 Å und 3.43 ± 0.91 Å bei pH 5.2 beziehungsweise 7.8, die sich nur in der katalytischen Triade unterscheiden. Bei tiefem pH sehen wir ein G_{syn}-AH^{+}_{anti} Basenpaar, während G4 bei hohem pH



Figur 6.5 Representative Lösungsstruktur von AvD5 bei pH 5.2. Einige Nukleotide, die im Text erwähnt werden, sind bezeichnet. Pentaloop: blau, Bulge: grün, katalytische Triade: rot.

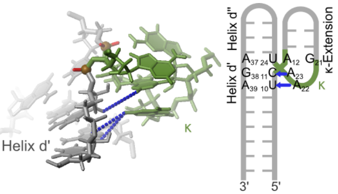
verschiedene Konformationen sowohl ins Innere der Helix als auch nach aussen gewandt, einnimmt. Obwohl eine kleine Anzahl an RDCs (residual dipolar couplings) in die Rechnungen miteinbezogen wurden, wirkt der Bulge immer noch als eine Art bewegliches Gelenk zwischen den beiden Helices, was zu vergleichsweise hohen Gesamt-RMSDs führt.

NMR-Experimente wurden nicht nur zur Strukturaufklärung von AvD5 sondern auch zur Identifizierung von Metallionenbindungsstellen mit möglichen strukturellen als auch katalytischen Funktionen angewandt. Die starke Mg(II)-Bindungsstelle, die wir in Fluoreszenzexperimenten bestimmen konnten, zeigt sich nur in einem anfänglichen Richtungswechsel der Änderung der chemischen Verschiebung einiger Resonanzen in der Nähe des Bulges (Figure 6.6). Die Veränderungen der ^{13}C and ^{15}N Verschiebungen lassen keine starke direkte Koordination an eine der Nukleobasen erkennen, was vermuten lässt, dass die Bindung am Zucker-Phosphat-Rückgrat des Bulges stattfindet. Zusätzlich zu dieser einen stärkeren Bindungsstelle zeigen die Mg(II) abhängigen Änderungen der ^{1}H chemischen Verschiebungen 4-5 weitere, schwächere Metallionenbindungsstellen an, darunter auch eine zweite in der Region des Bulges. Alle haben Affinitäten im niedrigen bis mittleren millimolaren Bereich, wobei der Pentaloop die geringste Affinität aufweist. Zusätzlich wurden Cd(II) und Kobalt(III)hexamin anstelle von Mg(II) eingesetzt um Informationen über die Art der Koordination — direkt oder indirekt, d.h. über die innere Schicht der Wasserliganden — zu gewinnen. Sowohl Cd(II)-abhängige Änderungen der chemischen Verschiebungen von ^{13}C und ^{15}N als auch intermolekulare NOEs zwischen RNA und Kobalt(III)hexamin zeigen eine bevorzugte Bindung an zwei Tandem-G-C-Basenpaaren in Helix 1 und Helix 2. Der Pentaloop bindet keines der getesteten Metallionen besonders fest.



Figur 6.6 Änderung der chemischen Verschiebung einer Resonanz in der Nähe des Bulges in Abhängigkeit von Mg(II), bestimmt aus ^{1}H,^{1}H-NOESY Spektren von AvD5 bei pD 6.8. Der Graph illustriert wie die Form der Kurve durch zwei Bindungsstellen im millimolaren und submillimolaren Bereich beschrieben werden kann. Die Anzahl der Datenpunkte reicht allerdings nicht für eine quantitative Analyse beider K_Ds aus.

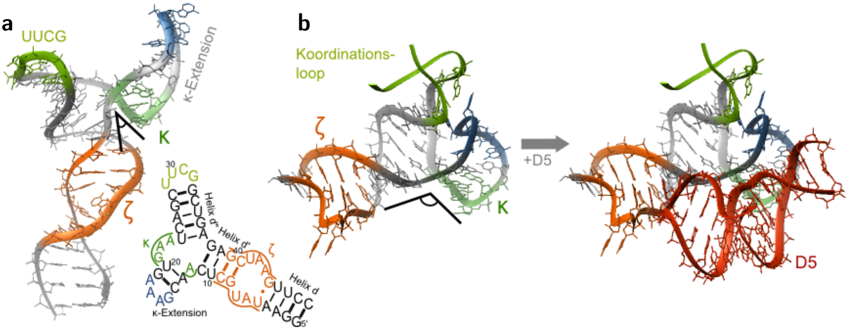
D1kz – die Struktur in Lösung und ihre Abhängigkeit von zwei- und dreiwertigen Metallionen



Figur 6.7 Die Wechselwirkung der κ-Adenine mit Helix d' in D1kz. Einige der beobachtbaren NOEs sind in blau eingezeichnet. Nukleotide die in der Struktur links gezeigt sind, sind rechts eingezeichnet, wo die Interaktion schematisch in den selben Farben dargestellt ist.

Unser 49-nt langes D1kz-Konstrukt umfasst die κ-ζ-Region von Sc.ai5γ. ζ ist ein 11-nt Tetraloop-Rezeptor und damit ein Motiv, das in anderem Zusammenhang schon in Lösung studiert worden ist[385,386] und von dem man weiss, dass es seine Konformation stark verändert, wenn ein Tetraloop bindet. In D1kz reagiert die Region empfindlich auf Metallionen, was sich im Verbreitern und Verschwinden der Peaks manifestiert. Zudem beobachten wir eine transiente intermolekulare Interaktion zwischen dem GAAA-Tetraloop eines D1kz-Moleküls und dem Rezeptor eines zweiten ab einer Mg(II)-Konzentration von ~4 mM, was die Entstehung von D1kz-Homodimeren nahelegt. Der Rezeptor bleibt allerdings in seiner "ungebundenen" Konformation, wie das Fehlen der charakteristischen Merkmale des Tetraloop-Rezeptor Komplexes in NMR-Spektren beweist. Eine kanonische Wechselwirkung des 11-nt-Rezeptors mit dem GAAA-Tetraloop ist vermutlich aus sterischen Gründen nicht möglich.

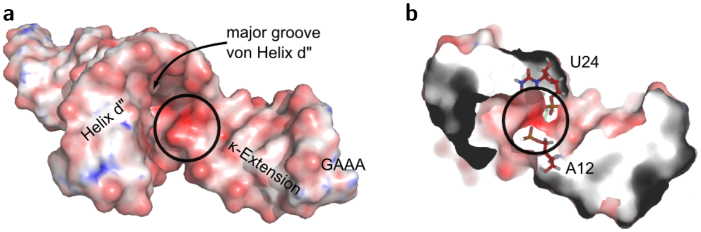
κ besteht aus vier Nukleotiden eingebettet in eine Dreierkreuzung ("three-way junction") — also dort wo drei RNS-Helices aufeinandertreffen — und bildet im Kontext des gefalteten Introns eine Struktur aus, die der eines GNRA-Tetraloops entspricht[237,266]. Unsere NMR Experimente zeigen, dass diese Struktur auch in Abwesenheit des Bindungspartners D5 und von zweiwertigen Metallionen ausgebildet wird. Zudem gibt es A-minor Interaktionen mit der benachbarten Helix d' (Figure 6.7). Letztere sind bemerkenswert, weil diese im gefalteten Intron, wo κ an κ' in D5 bindet[236,266] (Figure 6.8), nicht vorkommen. Zugabe von ScD5 (D5 von Sc.ai5γ) zu D1kz in Lösung ist nicht ausreichend um die A-minor Interaktion zu lösen. Sogar bei starkem ScD5-Überschuss und in Anwesenheit von Mg(II) gibt es kein Zeichen einer D5-D1kz Wechselwirkung. Wir vermuten daher, dass ein Netzwerk tertiärer Interaktionen im gefalteten Intron, möglicherweise unter Einbezug des benachbarten sogenannten "Koordinationsloops" (Figure 6.8b), benötigt wird um sowohl die κ-ζ-Region als auch D5 in ihre jeweiligen bindungs-kompetenten Formen zu bringen.



Figur 6.8 Die unterschiedliche Orientierung der κ-Extension in Lösung (in D1kz) (a) und in der κ-ζ-Region in der Gruppe IIC Intron-Kristallstruktur (b) (PDB-ID: 3G78[237]). In (a) verbinden A-minor-Wechselwirkungen κ mit Helix d', sodass der Winkel zwischen den beiden klein ist. Werden diese Interaktionen unterbrochen (b) entsteht Platz und D5 kann andocken. Letzeres wird möglicherweise durch die Wechselwirkung zwischen der κ-Extension und dem benachbarten "Koordinationsloop" erleichtert.

Während Mg(II) bei der Ausbildung der beobachteten A-minor Interaktionen keine Rolle spielt, stellte es sich als essentieller stabilisierender Faktor für die Basenpaare in der Umgebung der κ-Dreierkreuzung heraus. In Lösungen welche nur einwertige Salze enthalten sind die direkt benachbarten Basenpaare instabil, was unter anderem am Fehlen ihrer Imino-Proton-Resonanzen erkennbar ist. Zusätzlich gibt es dort Lücken in der sequentiellen Zuordnung ("sequential walk") der Resonanzen, wo die beiden Stränge, die in die κ-Extension hinein- und wieder herausführen, sich kreuzen (C11, A12, A37, G38). Wird Mg(II) zugegeben werden die Basenpaare stabilisiert und die fehlenden Resonanzen tauchen auf. Auch ein Cross-Peak zwischen U24H3 und G38H1 erscheint, der Helix d' und d" über die Dreierkreuzung hinweg verbindet, ein deutlicher Hinweis auf co-axiales Stacking der beiden Helices, was auch in den vorhandenen Kristallstrukturen beobachtet wird[236,280]. Die Mg(II)-abhängige Stabilisierung der κ-Dreierkreuzung führt die Struktur also in Richtung der Konformation, die wir auch im Intron finden und weckt Assoziationen zur Metallionenabhängigkeit der anfänglichen Kompaktierung von D1, bei der eine Stabilisierung der Helices in der κ-ζ-Region auch eine Rolle spielt.

^{1}H,^{1}H-NOESY-Spektren von D1kz in 10 mM $MgCl_2$ lieferten Distanzinformation, die in Molekulardynamiksimulationen zur Bestimmung eines Strukturensembles mit einem RMSD von 4.30 ± 1.23 Å verwendet wurden (Figure 6.8a). Visualisierung des Oberflächenpotentials an der Dreierkreuzung lässt eine Region starken negativen Potentials über den Phosphaten von A12 und U24 erkennen (Figure 6.9), die Metallionen anziehen könnte. Die Stabilisierung der κ-Region geht mit starken Änderungen in den chemischen Verschiebungen von ^{1}H Resonanzen einher. Die Überlappung von strukturellen und Metallbindungseinflüssen macht es jedoch schwierig, die Bindungsstelle genau zu lokalisieren. Chemische Verschiebungen von ^{13}C und ^{15}N Resonanzen lassen keine direkte Koordination an den Nukleobasen erkennen, was für Mg(II)-Bindung am Phosphatrückgrat spricht.



Figur 6.9 a) Eine Region hohen elektrostatischen Oberflächenpotentials zwischen Helix d" und der κ-Extension bezeichnet die Lage der Phosphatgruppen von A12 und U24 (b). Das Potential ist durch einen Farbgradienten von rot (-1068 mV) nach blau (134 mV) dargestellt.

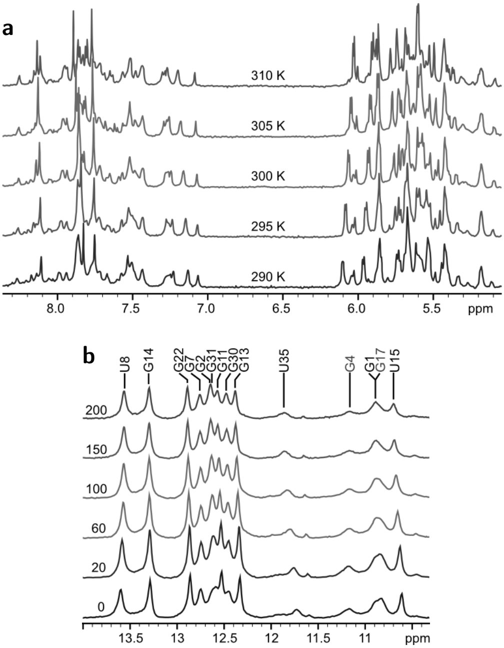
Die beobachtete Stabilisierung in der κ-Region ist nicht Mg(II)-spezifisch, Cd(II) und Kobalt(III)hexamin sind darin sogar noch effizienter, d.h. sie sind schon bei niedrigeren Konzentrationen wirksam. Chemische Verschiebungen von ^{13}C und ^{15}N und intermolekulare NOEs weisen auf eine Wechselwirkung beider Metallionen mit der major groove von Helix d" hin. Da nicht erwartet wird, dass Cd(II) und Kobalt(III)hexamin auf dieselbe Weise binden, führen wir die stabilisierende Wirkung auf die Ladungsneutralisierung in einer Region mit konzentriertem elektronegativem Potential zwischen Helix d" und κ-Extension zurück, die keine spezifische Koordinationsweise voraussetzt (Figure 6.9).

Konklusion

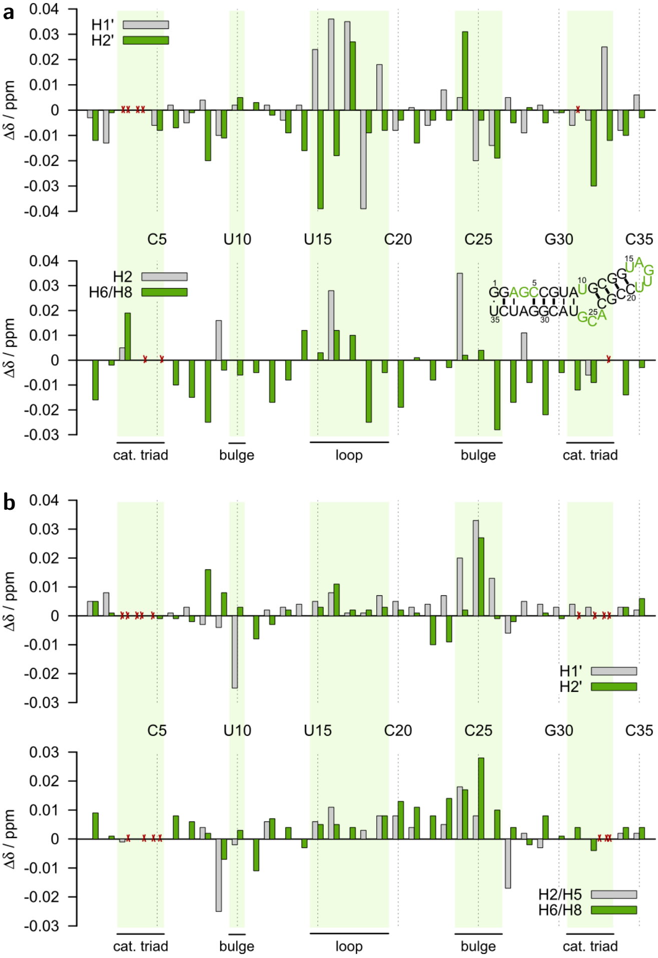
Die zwei in dieser Arbeit untersuchten RNS-Konstrukte sind nur kleine Teile der riesigen Gruppe II Intron Ribozyme aus denen sie stammen. Für beide zeigt sich klar, dass die Strukturen in Lösung sich von denen in der Kristallstruktur und wohl auch im komplett gefalteten Intron unterscheiden. Besonders die nicht-kanonischen, ungepaarten Regionen sind durch diverse Grade an Flexibilität gekennzeichnet. Nichtsdestotrotz wird der natürliche Kofaktor Mg(II), der sowohl für die Faltung als auch für die katalytische Aktivität von Gruppe II Introns unabkömmlich ist, in Lösung von den gleichen Stellen angezogen, an denen er auch in der vollen Intronstruktur zu finden ist. Mg(II) beeinflusst Konformation und Flexibilität von spezifischen Teilstrukturen und führt dabei sowohl den AvD5 Bulge als auch die Region um die D1kz-Dreierkreuzung in Richtung der Konformation, die man im gefalteten Gruppe II Intron findet. Unsere Ergebnisse weisen darauf hin, dass sowohl die Anpassungsfähigkeit der flexiblen, nicht-kanonischen Regionen als auch die Vorauswahl von bestimmten Konformationen durch Mg(II) eine Rolle bei der Assemblierung grosser RNS-Moleküle spielen. Auf diese Weise kann das, was wir in unseren beiden kleinen RNS-Konstrukten beobachten, zum Verständnis der grundlegenden Prinzipien der RNS-Faltung beitragen.

Appendices

Appendix 1. 1D ^{1}H spectra of AvD5 at increasing temperature and KCl concentration. a) Temperature dependence of aromatic protons and anomeric sugar protons in AvD5 (0.35 mM AvD5, pD 6.8, 60 mM KCl, 10 µM KCl). b) 1D spectra of imino protons at increasing KCl concentrations (0.7 mM AvD5, pH 7.2, 275 K, 10 µM EDTA). Tentatively assigned G4 and G17 are marked in red



Appendix 2. Temperature and monovalent salt dependent chemical shift changes in AvD5. a) Chemical shift changes of AvD5 protons when the temperature is raised from 290 K to 300 K. b) Chemical shift changes of AvD5 protons when raising the KCl concentration from 60 to 120 mM. To distinguish resonances that are not moving at all from values that are missing because assignment is not possible under all conditions the latter were crossed out red in the graphs. The secondary structure scheme indicates catalytic triad, bulge and loop in green.



Appendix 3. AvD5 ^{1}H chemical shifts at pH 6.7. Non-exchangeable proton resonances were recorded at 300 K, exchangeable ones at 280 K. RNA concentrations 0.4-1 mM, 60 mM KCl, 10 µM EDTA, 600 and 700 MHz. ^{a}Ade-H2/Pyr-H5, ^{b}Pyr-H6/Pur-H8, ^{c}Gua-H21/Cyt-H41/Ade-H61, ^{d}Gua-H22/Cyt-H42/Ade-H62, ^{e}Ura-H3/Gua-H1.

	H1'	H2'	H3'	H4'	H5'	H5"	H2/H5^a	H6/H8^b	amino-H1^c	amino-H2^d	imino-H^e
G1	5.876	4.935	4.583	4.559	4.41	4.284	-	8.14	-	-	10.864
G2	5.78	4.683	4.626	4.557	-	4.243	-	7.653	-	-	12.482
A3	-	4.327	-	-	-	-	7.595	-	-	-	-
G4	-	-	-	-	-	-	-	-	-	-	11.099
C5	5.58	-	-	-	-	-	5.176	7.523	8.454	6.884	-
C6	5.432	4.545	4.386	4.399	4.526	4.084	5.417	7.74	8.283	6.763	-
G7	5.668	4.479	4.507	4.446	-	4.088	-	7.501	-	6.001	12.779
U8	5.564	4.541	4.585	4.456	-	4.119	5.097	7.71	-	-	13.545
A9	6.019	4.533	4.562	4.527	4.557	4.179	7.334	8.099	-	-	-
U10	5.479	4.293	4.588	4.403	-	4.106	5.393	7.544	-	-	-
G11	5.733	4.578	4.521	4.476	4.419	4.162	-	7.829	-	-	12.544
C12	5.476	4.495	4.509	4.417	4.485	4.116	5.187	7.645	8.405	6.697	-
G13	5.706	4.596	4.562	4.439	-	4.105	-	7.521	8.013	-	12.366
G14	5.673	4.523	4.276	4.438	4.431	4.034	-	7.072	8.715	5.995	13.302
U15	5.473	4.366	4.449	4.354	4.357	4.047	5.353	7.51	-	-	10.635
A16	5.704	4.591	4.56	4.391	4.296	4.047	7.856	8.12	-	-	-
G17	5.568	4.595	4.682	4.436	4.192	4.044	-	7.759	-	-	10.828
U18	6.063	4.445	-	4.551	4.214	4.168	5.931	7.829	-	-	-
U19	5.763	4.469	4.545	4.533	4.303	4.252	5.859	7.855	-	-	-
C20	5.605	4.449	4.485	4.471	4.53	4.215	5.868	7.963	8.423	7.253	-
C21	5.52	4.587	4.466	4.467	4.496	4.158	5.598	7.717	8.459	6.783	-
G22	5.69	4.499	4.51	4.458	4.473	4.103	-	7.573	8.173	-	12.873
C23	5.425	4.387	4.422	-	-	4.062	5.177	7.509	8.19	6.88	-
A24	5.879	4.249	4.532	4.42	-	4.074	7.195	7.934	-	-	-
C25	5.576	4.114	4.488	4.364	4.522	4.072	5.418	7.47	-	-	-
G26	5.624	4.696	4.585	4.456	4.263	4.145	-	7.842	-	-	-
U27	5.702	4.516	4.612	4.445	-	4.202	5.582	7.851	-	-	-
A28	6.009	4.58	4.707	4.535	4.552	4.219	7.271	8.267	7.848	6.435	-
C29	5.436	4.433	4.415	4.428	4.499	4.099	5.178	7.474	8.133	6.757	-
G30	5.691	4.611	4.494	4.475	4.475	4.101	-	7.449	-	-	12.481
G31	5.688	-	4.603	4.44	4.511	4.1	-	7.261	-	-	12.649
A32	6.057	-	-	-	-	-	-	7.978	-	-	-
U33	-	-	-	-	-	-	-	-	-	-	-

	H1'	H2'	H3'	H4'	H5'	H5"	H2/H5[a]	H6/H8[b]	amino-H1[c]	amino-H2[d]	imino-H[e]
C34	5.661	4.339	4.352	-	-	4.011	5.578	7.808	8.371	7.061	-
U35	5.725	4.004	4.17	4.163	4.527	-	5.625	7.774	-	-	11.945

Appendix 4. AvD5 ^{13}C and ^{15}N chemical shifts at pH 6.7. ^{13}C shifts were recorded at 290 K, ^{15}N shifts at 280 K. RNA concentrations 0.4-1 mM, 60 mM KCl, 10 µM EDTA, 600 and 700 MHz. [a]Ade-C2/Pyr-C5, [b]Pyr-C6/Pur-C8.

	C1'	C2'	C3'	C4'	C5'	C2/C5[a]	C6/C8[b]	N1	N3	N7
	C1'	C2'	C3'	C4'	C5'	C2/C5	C6/C8	N1	N3	N7
G1	91.929	75.459	72.977	-	67.356	-	139.67	-	-	233.286
G2	-	-	-	-	-	-	137.308	146.822	-	234.198
A3	-	-	-	-	-	153.868	139.599	-	-	-
G4	-	-	-	-	-	-	-	-	-	-
C5	-	75.438	-	-	-	-	-	-	197.887	-
C6	93.94	75.363	-	81.894	64.5	97.638	140.863	-	195.951	-
G7	93.034	-	-	-	-	-	136.263	147.18	-	235.183
U8	93.314	-	-	-	64.568	102.854	141.5	-	161.711	-
A9	92.928	-	-	-	65.413	153.263	139.936	-	-	232.122
U10	92.529	75.609	73.21	82.879	64.904	104.298	141.004	-	-	-
G11	92.883	-	-	-	66.116	-	137.551	147.832	-	232.687
C12	93.599	75.471	-	-	64.665	97.444	140.886	-	196.973	-
G13	92.721	-	-	-	-	-	136.171	146.683	-	235.001
G14	92.832	-	73.042	-	65.657	-	135.714	148.517	-	236.279
U15	93.291	75.412	73.041	82.971	64.802	104.195	140.576	-	157.404	-
A16	91.776	75.467	73.809	83.705	66.362	155.299	140.774	225.185	-	231.661
G17	89.631	76.266	78.288	85.705	67.71	-	139.86	-	-	236.088
U18	90.127	75.951	77.927	85.296	68.299	106.003	143.775	-	-	-
U19	92.809	75.128	72.517	83.888	68.135	104.508	143.849	-	-	-
C20	93.668	75.624	-	-	66.123	98.1	142.456	-	197.597	-
C21	94.09	75.385	-	-	65.698	98.371	141.144	-	196.792	-
G22	92.693	-	-	-	-	-	136.075	147.543	-	234.59
C23	93.399	75.421	72.381	-	64.612	97.091	140.392	-	196.34	-
A24	92.213	76.241	73.33	82.198	65.402	153.926	139.648	-	-	-
C25	91.616	75.847	72.284	-	66.434	98.14	141.944	-	-	-
G26	91.066	75.138	-	-	-	-	139.071	-	-	235.699
U27	92.79	-	-	-	65.874	104.147	142.304	-	-	-

	C1’	C2’	C3’	C4’	C5’	C2/C5^a	C6/C8^b	N1	N3	N7
A28	92.724	75.501	73.286	82.246	65.221	153.461	140.173	222.176	-	229.165
C29	93.641	75.601	72.262	-	64.496	97.232	140.392	-	196.84	-
G30	92.873	-	-	-	-	-	136.069	146.72	-	234.916
G31	92.799	-	-	82.555	-	-	136.498	147.81	-	233.898
A32	-	-	-	-	-	154.612	-	-	-	-
U33	94.106	-	-	-	-	104.186	-	-	-	-
C34	93.084	75.322	-	-	64.534	97.421	141.613	-	195.959	-
U35	-	-	-	-	-	104.633	142.147	-	-	-

Appendix 5. AvD5 ^{1}H chemical shifts at pH 5.2. Non-exchangeable proton resonances were recorded at 300 K, exchangeable ones at 275 K. RNA concentrations 0.4-1 mM, 60 mM KCl, 10 µM EDTA, 600 and 700 MHz. [a]Ade-H2/Pyr-H5, [b]Pyr-H6/Pur-H8, [c]Gua-H21/Cyt-H41/Ade-H61, [d]Gua-H22/Cyt-H42/Ade-H62, [e]Ura-H3/Gua-H1.

	H1’	H2’	H2/H5[a]	H6/H8[b]	amino-H1[c]	amino-H2[d]	imino-H[e]
G1	5.865	4.921	-	8.114	-	-	10.85
G2	5.756	4.676	-	7.655	-	-	12.564
A3	6.005	4.99	7.588	7.922	7.807	6.881	-
G4	5.943	4.633	-	7.723	-	-	11.136
C5	5.668	4.308	5.144	7.75	8.446	7.043	-
C6	5.437	4.55	5.404	7.689	8.288	6.871	-
G7	5.675	4.495	-	7.565	-	-	12.884
U8	-	-	5.087	7.663	-	-	13.582
A9	6.024	4.649	-	-	-	-	-
U10	5.353	4.384	5.393	7.603	-	-	-
G11	5.737	4.537	-	7.736	-	-	12.492
C12	5.453	4.456	5.21	7.625	8.386	6.823	-
G13	5.688	4.579	-	7.607	-	-	12.386
G14	5.666	4.511	-	7.155	8.706	-	13.272
U15	5.582	4.376	5.456	7.585	-	-	10.708
A16	5.752	4.589	7.964	8.21	-	-	-
G17	5.596	4.666	-	7.747	-	-	10.843
U18	6.021	4.431	5.906	7.808	-	-	-
U19	5.782	4.473	5.872	7.865	-	-	-
C20	5.596	4.457	5.881	7.96	8.454	7.282	-
C21	5.516	4.454	-	7.748	8.466	6.85	-
G22	-	-	-	-	-	-	12.827
C23	-	-	-	-	8.335	6.924	-
A24	-	-	-	-	-	-	-
C25	-	-	-	-	-	-	-
G26	-	-	-	-	-	-	-
U27	-	4.473	-	-	-	-	-
A28	5.915	4.603	-	8.284	-	-	-
C29	5.476	4.392	5.202	7.494	8.155	6.849	-
G30	5.72	4.644	-	7.586	-	-	12.272
G31	5.721	4.611	-	7.286	-	-	12.385
A32	6.024	4.526	8.216	8.001	10.387	9.206	-
U33	5.596	4.358	5.147	7.773	-	-	14.464
C34	5.642	4.312	5.573	7.796	8.455	7.135	-
U35	5.701	3.982	5.589	7.764	-	-	11.773

Appendix 6. AvD5 ^{13}C and ^{15}N chemical shifts at pH 5.2. AvD5 ^{13}C and ^{15}N chemical shifts at pH 5.2. ^{13}C shifts were recorded at 300 K, ^{15}N shifts at 275 K. RNA concentrations 0.4-1 mM, 60 mM KCl, 10 µM EDTA, 600 and 700 MHz. ^{a}Pyr-C6/Pur-C8, ^{b}Cyt-N4/Ade-N6.

	C2	C6/C8^{a}	N1	N3	N4/N6^{b}
G1	-	139.523	144.76	-	-
G2	-	137.148	146.776	-	-
A3	153.757	139.586	222.001	-	83.553
G4	-	138.995	145.662	-	-
C5	-	141.524	-	196.449	99.24
C6	-	140.855	-	196.514	97.811
G7	-	136.512	147.245	-	-
U8	-	141.907	-	162.089	-
A9	-	-	-	-	-
U10	-	140.829	-	-	-
G11	-	136.62	147.538	-	-
C12	-	140.629	-	-	98.293
G13	-	136.165	146.853	-	-
G14	-	135.973	148.476	-	-
U15	-	141.408	-	157.463	-
A16	152.143	142.236	-	-	-
G17	-	140.231	146.929	-	-
U18	-	143.906	-	-	-
U19	-	143.922	-	-	-
C20	-	142.692	-	-	98.961
C21	-	141.191	-	196.653	97.551
G22	-	136.228	147.612	-	-
C23	-	-	-	-	98.372
A24	-	-	-	-	-
C25	-	-	-	-	-
G26	-	139.074	-	-	-
U27	-	-	-	-	-
A28	-	140.828	-	-	-
C29	-	140.461	-	-	98.375
G30	-	-	146.696	-	-
G31	-	136.562	146.789	-	-
A32	147.13	142.495	-	-	95.78
U33	-	142.365	-	162.879	-
C34	-	141.673	-	196.526	99.366
U35	-	-	-	158.144	-

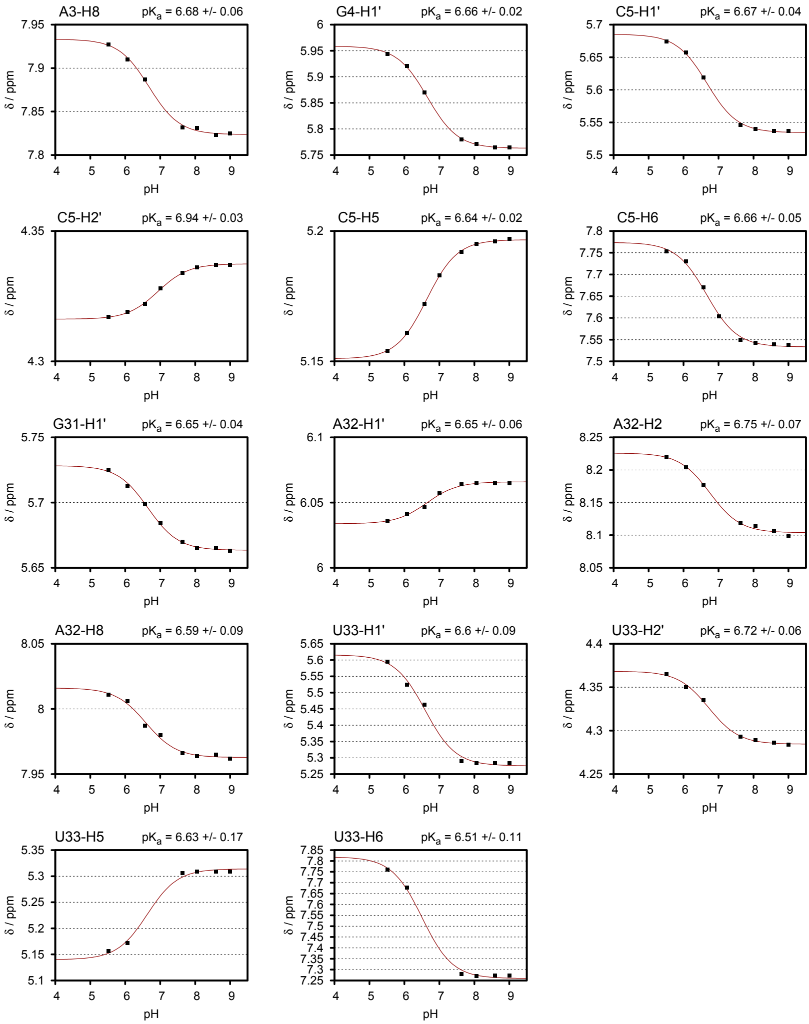
Appendix 7. AvD5 ^{1}H chemical shifts at pH 7.8. AvD5 ^{1}H chemical shifts at pH 7.8. Non-exchangeable proton resonances were recorded at 300 K, exchangeable ones at 275 K. RNA concentrations 0.4-1 mM, 60 mM KCl, 10 µM EDTA, 600 and 700 MHz. ^{a}Ade-H2/ Pyr-H5, ^{b}Pyr-H6/Pur-H8, ^{c}Gua-H21/Cyt-H41/Ade-H61, ^{d}Gua-H22/Cyt-H42/Ade-H62, ^{e}Ura-H3/Gua-H1.

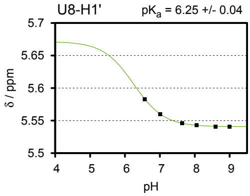
	H1'	H2'	H2/H5^{a}	H6/H8^{b}	amino-H1^{c}	amino-H2^{d}	imino-H^{e}
G1	5.849	4.908	-	8.157	-	-	10.91
G2	5.761	4.674	-	7.638	-	-	12.608
A3	5.927	-	7.588	7.813	-	-	-
G4	5.759	-	-	7.684	-	-	11.165
C5	5.534	4.326	5.186	7.539	8.498	6.912	-
C6	5.419	4.53	5.412	7.746	8.306	6.812	12.735
G7	5.657	4.467	-	7.478	8.145	6.016	-
U8	5.536	4.537	5.087	7.705	-	-	13.575
A9	6.009	4.504	7.283	8.049	-	-	-
U10	5.543	4.254	5.387	7.52	-	-	11.527
G11	5.723	4.605	-	7.861	8.149	6.029	12.542
C12	5.478	4.489	5.177	7.637	8.413	6.734	-
G13	5.703	4.594	-	7.517	7.983	5.832	12.345
G14	5.664	4.52	-	7.062	8.724	6.004	13.294
U15	5.452	4.356	5.332	7.497	-	-	10.66
A16	5.687	4.579	7.839	8.106	-	-	-
G17	5.557	4.585	-	7.749	-	-	10.878
U18	6.058	4.437	5.924	7.819	-	-	-
U19	5.743	4.455	5.844	7.843	-	-	-
C20	5.596	4.439	5.854	7.95	8.462	7.279	-
C21	5.509	4.581	5.588	7.705	8.491	6.804	-
G22	5.677	4.495	-	7.554	8.207	5.846	12.87
C23	5.402	4.381	5.171	7.465	8.188	6.919	-
A24	5.848	4.193	7.146	7.919	-	-	-
C25	5.551	4.083	5.381	7.403	-	-	-
G26	5.627	4.711	-	7.836	-	-	11.113
U27	5.665	4.519	5.582	7.846	-	-	-
A28	6.021	4.567	7.233	8.254	7.89	6.522	-
C29	5.411	4.434	5.173	7.454	8.157	6.845	-
G30	5.668	4.592	-	7.425	7.974	5.768	12.447
G31	5.66	4.312	-	7.255	-	-	12.619
A32	6.054	4.662	8.098	7.956	-	-	-
U33	5.28	4.283	5.296	7.266	-	-	-
C34	5.652	4.333	5.573	7.787	8.472	7.105	-
U35	5.721	3.999	5.623	7.76	-	-	11.933

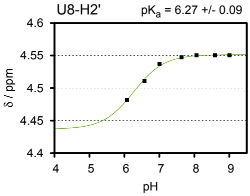
Appendix 8. AvD5 ^{13}C and ^{15}N chemical shifts at pH 7.8. AvD5 ^{13}C and ^{15}N chemical shifts at pH 7.8. ^{13}C shifts were recorded at 300 K, ^{15}N shifts at 275 K. RNA concentrations 0.4-1 mM, 60 mM KCl, 10 µM EDTA, 600 and 700 MHz. ^{a}Pyr-C6/Pur-C8, ^{b}Cyt-N4/Ade-N6.

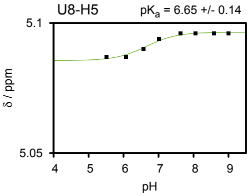
	C2	C6/C8^{a}	N1	N3	N4/N6^{b}
G1	-	139.781	143.336	-	-
G2	-	137.42	146.888	-	-
A3	153.861	139.555	-	-	-
G4	-	-	145.936	-	-
C5	-	-	-	197.671	97.659
C6	-	140.892	-	195.837	97.992
G7	-	136.225	147.159	-	-
U8	-	141.451	-	161.746	-
A9	153.789	139.688	-	-	-
U10	-	141.153	-	-	-
G11	-	137.797	147.76	-	-
C12	-	140.924	-	196.8	97.794
G13	-	136.182	146.657	-	-
G14	-	135.674	148.488	-	-
U15	-	140.479	-	157.347	-
A16	155.363	140.7	-	-	-
G17	-	139.82	147.098	-	-
U18	-	143.752	-	-	-
U19	-	143.801	-	-	-
C20	-	142.447	-	197.289	99.114
C21	-	141.088	-	196.627	97.594
G22	-	136.021	147.489	-	-
C23	-	140.349	-	196.067	98.221
A24	154.033	139.584	-	-	-
C25	-	141.839	-	-	-
G26	-	139.066	-	-	-
U27	-	142.162	-	-	-
A28	153.394	140.061	222.078	-	84.597
C29	-	140.344	-	196.654	98.291
G30	-	136.026	146.721	-	-
G31	-	136.416	147.755	-	-
A32	154.619	139.689	-	-	-
U33	-	141.056	-	-	-
C34	-	141.585	-	-	99.181
U35	-	142.11	-	157.883	-

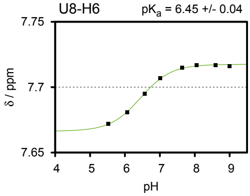
Appendix 9. Least-squares fits of pH dependent chemical shift changes yielding pK_a values in the catalytic triad and bulge of AvD5. The red (catalytic triad) and green (bulge) curves are fits to Equation 4.5 (see Experimental Section).

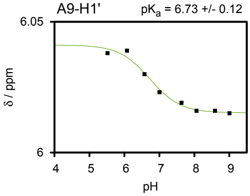


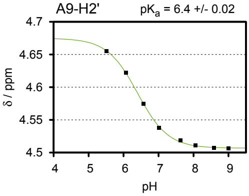

U8-H1' pK$_a$ = 6.25 +/- 0.04

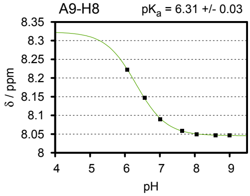

U8-H2' pK$_a$ = 6.27 +/- 0.09

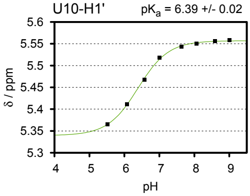

U8-H5 pK$_a$ = 6.65 +/- 0.14

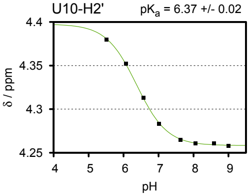

U8-H6 pK$_a$ = 6.45 +/- 0.04

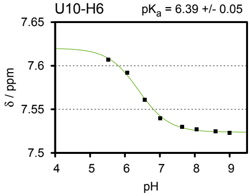

A9-H1' pK$_a$ = 6.73 +/- 0.12

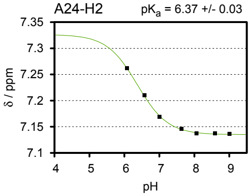

A9-H2' pK$_a$ = 6.4 +/- 0.02

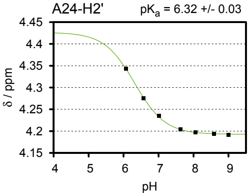

A9-H8 pK$_a$ = 6.31 +/- 0.03

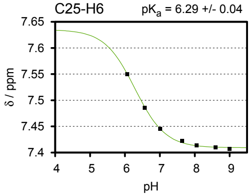

U10-H1' pK$_a$ = 6.39 +/- 0.02

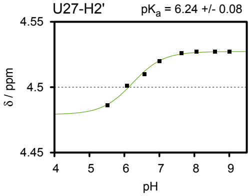

U10-H2' pK$_a$ = 6.37 +/- 0.02

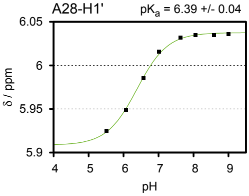

U10-H6 pK$_a$ = 6.39 +/- 0.05

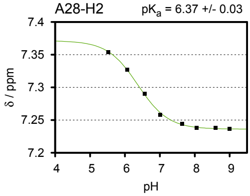

A24-H2 pK$_a$ = 6.37 +/- 0.03

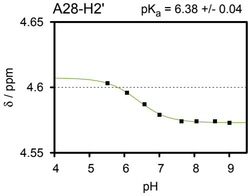

A24-H2' pK$_a$ = 6.32 +/- 0.03

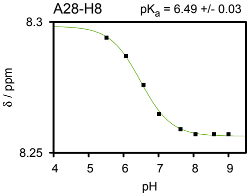

C25-H6 pK$_a$ = 6.29 +/- 0.04


U27-H2' pK$_a$ = 6.24 +/- 0.08


A28-H1' pK$_a$ = 6.39 +/- 0.04


A28-H2 pK$_a$ = 6.37 +/- 0.03


A28-H2' pK$_a$ = 6.38 +/- 0.04


A28-H8 pK$_a$ = 6.49 +/- 0.03

Appendix 10. NOE-derived distance restraints used in AvD5 structure calculations. Distance restraints derived from ^{1}H,^{1}H-NOESY spectra recorded in D_2O at 300 K (non-exchangeable protons) and in H_2O at 275 K (exchangeable protons), pD/pH 6.7, 60 mM KCl, 10 µM EDTA. Restraints for the catalytic triad region are listed separately below. There are two sets of restraints for the catalytic triad for low and high pH that were used in two independent sets of calculations. This was not necessary for restraints in other regions — because they do not change significantly with pH or are not observable at low pH (see bulge) — and the same restraints were used in both the high and low pH set of calculations. Restraints are given in the following format (distances in Å):

<first proton> <second proton> <target distance> <lower limit> <upper limit>

Non-exchangeable protons:

G1

1-H1'	1-H8	3.8	2.0	0.7
1-H2'	1-H8	3.8	2.0	0.7
1-H3'	1-H8	3.8	2.0	0.7
1-H4'	1-H8	5.0	2.0	1.0
1-H5'	1-H8	3.8	2.0	0.7
1-H5"	1-H8	3.8	2.0	0.7
1-H1'	2-H8	5.0	2.0	1.0
1-H2'	2-H8	3.8	2.0	0.7
1-H3'	2-H8	3.8	2.0	0.7
1-H8	2-H8	5.0	2.0	1.0
1-H1'	2-H1'	5.0	2.0	1.0
1-H2'	2-H1'	6.0	2.0	1.0

G2

2-H1'	2-H8	3.8	2.0	0.7
2-H2'	2-H8	5.0	2.0	1.0
2-H3'	2-H8	3.8	2.0	0.7
2-H5'	2-H8	3.8	2.0	0.7
2-H5"	2-H8	3.8	2.0	0.7
2-H1'	3-H8	5.0	3.0	1.0
2-H2'	3-H8	3.8	2.0	0.7
2-H8	3-H8	5.0	2.0	1.0

C6

6-H1'	6-H6	3.8	2.0	0.7
6-H2'	6-H6	3.8	2.0	0.7
6-H3'	6-H6	3.8	2.0	0.7
6-H4'	6-H6	5.0	2.0	1.0
6-H5'	6-H6	5.0	2.0	1.0
6-H5"	6-H6	5.0	2.0	1.0
6-H2'	6-H5	5.0	2.0	1.0
6-H3'	6-H5	5.0	2.0	1.0
6-H1'	7-H8	5.0	3.0	1.0
6-H2'	7-H8	2.5	0.7	0.5
6-H3'	7-H8	3.8	2.0	0.7
6-H6	7-H8	5.0	2.0	1.0
6-H1'	7-H1'	5.0	2.0	1.0
6-H2'	7-H1'	5.0	3.0	1.0

G7

7-H1'	7-H8	3.8	2.0	0.7
7-H2'	7-H8	3.8	2.0	0.7
7-H3'	7-H8	3.8	2.0	0.7
7-H4'	7-H8	5.0	2.0	1.0
7-H5'	7-H8	5.0	2.0	1.0
7-H5"	7-H8	3.8	2.0	0.7
7-H1'	8-H6	5.0	2.0	1.0
7-H2'	8-H6	2.5	0.7	0.5
7-H3'	8-H6	3.8	2.0	0.7
7-H8	8-H6	5.0	2.0	1.0
7-H1'	8-H5	5.0	2.0	1.0
7-H2'	8-H5	3.8	2.0	0.7
7-H3'	8-H5	5.0	2.0	1.0
7-H8	8-H5	5.0	2.0	1.0
7-H2'	8-H2'	5.0	2.0	1.0
7-H1'	8-H1'	5.0	2.0	1.0
7-H2'	8-H1'	3.8	2.0	0.7

U8

8-H1'	8-H6	3.8	2.0	0.7
8-H2'	8-H6	3.8	2.0	0.7
8-H3'	8-H6	3.8	2.0	0.7
8-H4'	8-H6	5.0	2.0	1.0
8-H5'	8-H6	3.8	2.0	0.7
8-H5"	8-H6	3.8	2.0	0.7
8-H1'	8-H5	5.0	2.0	1.0
8-H2'	8-H5	5.0	2.0	1.0
8-H3'	8-H5	5.0	2.0	1.0
8-H5"	8-H5	6.0	2.0	1.0
8-H1'	9-H8	5.0	2.0	1.0
8-H2'	9-H8	2.5	0.7	0.5
8-H3'	9-H8	3.8	2.0	1.0
8-H6	9-H8	5.0	2.0	1.0
8-H1'	9-H1'	5.0	2.0	1.0
8-H2'	9-H1'	5.0	2.0	1.0
8-H1'	28-H2	6.0	3.0	1.0
8-H2'	28-H2	6.0	2.0	1.0

A9

9-H1'	9-H8	5.0	2.0	1.0
9-H2'	9-H8	5.0	2.0	1.0
9-H3'	9-H8	3.8	2.0	0.7
9-H4'	9-H8	5.0	2.0	1.0
9-H5'	9-H8	5.0	2.0	1.0
9-H5"	9-H8	5.0	2.0	1.0
9-H1'	9-H2	5.0	2.0	1.0
9-H2'	9-H2	5.0	2.0	1.0
9-H1'	10-H6	5.0	2.0	1.0
9-H2	10-H6	6.0	2.0	1.0
9-H2'	10-H6	2.5	0.7	0.5
9-H3'	10-H6	3.8	2.0	0.7
9-H8	10-H6	5.0	2.0	1.0
9-H1'	10-H5"	6.0	2.0	1.0
9-H1'	10-H5	5.0	2.0	1.0
9-H2'	10-H5	5.0	2.0	1.0
9-H3'	10-H5	5.0	2.0	1.0
9-H8	10-H5	5.0	2.0	1.0
9-H1'	10-H4'	6.0	2.0	1.0
9-H2	10-H2'	6.0	2.0	1.0
9-H2'	10-H2'	5.0	2.0	1.0
9-H1'	10-H1'	5.0	2.0	1.0
9-H2	10-H1'	3.8	2.0	0.7
9-H2'	10-H1'	5.0	2.0	1.0
9-H2	28-H8	6.0	2.0	1.0
9-H2	28-H2'	5.0	2.0	1.0
9-H1'	28-H2	3.8	2.0	0.7
9-H2	28-H2	5.0	3.0	1.0
9-H2'	28-H2	5.0	2.0	1.0
9-H8	28-H2	6.0	3.0	1.0
9-H2	28-H1'	3.8	2.0	0.7

U10

10-H1'	10-H6	3.8	2.0	0.7
10-H2'	10-H6	3.8	2.0	0.7
10-H3'	10-H6	3.8	2.0	0.7
10-H4'	10-H6	5.0	2.0	1.0
10-H5'	10-H6	5.0	2.0	1.0
10-H5"	10-H6	5.0	2.0	1.0
10-H1'	10-H5	5.0	2.0	1.0
10-H2'	10-H5	5.0	2.0	1.0
10-H3'	10-H5	5.0	2.0	1.0
10-H1'	11-H8	5.0	2.0	1.0
10-H2'	11-H8	2.5	0.7	0.5
10-H6	11-H8	5.0	2.0	1.0
10-H2'	11-H2'	6.0	2.0	1.0
10-H1'	11-H1'	6.0	2.0	1.0
10-H2'	11-H1'	6.0	3.0	1.0

G11

11-H1'	11-H8	3.8	2.0	0.7
11-H2'	11-H8	3.8	2.0	0.7
11-H3'	11-H8	3.8	2.0	0.7
11-H4'	11-H8	5.0	2.0	1.0

11-H5’	11-H8	5.0	2.0	1.0
11-H5"	11-H8	3.8	2.0	0.7
11-H1’	12-H6	5.0	2.0	1.0
11-H2’	12-H6	2.5	0.7	0.5
11-H3’	12-H6	3.8	2.0	0.7
11-H8	12-H6	5.0	2.0	1.0
11-H1’	12-H5	5.0	2.0	1.0
11-H2’	12-H5	3.8	2.0	0.7
11-H3’	12-H5	5.0	2.0	1.0
11-H8	12-H5	5.0	2.0	1.0
11-H2’	12-H2’	5.0	2.0	1.0
11-H1’	12-H1’	5.0	2.0	1.0
11-H2’	12-H1’	5.0	3.0	1.0
11-H1’	24-H2	3.8	2.0	0.7
11-H2’	24-H2	6.0	2.0	1.0
11-H8	24-H2	5.0	2.0	1.0
C12				
12-H1’	12-H6	3.8	2.0	0.7
12-H2’	12-H6	3.8	2.0	0.7
12-H3’	12-H6	3.8	2.0	0.7
12-H4’	12-H6	5.0	2.0	1.0
12-H5"	12-H6	3.8	2.0	0.7
12-H1’	12-H5	5.0	2.0	1.0
12-H2’	12-H5	5.0	2.0	1.0
12-H3’	12-H5	5.0	2.0	1.0
12-H5"	12-H5	5.0	2.0	1.0
12-H1’	13-H8	4.5	2.0	0.7
12-H2’	13-H8	2.5	0.7	0.5
12-H3’	13-H8	3.8	2.0	0.7
12-H6	13-H8	5.0	2.0	1.0
12-H2’	13-H2’	5.0	2.0	1.0
12-H1’	13-H1’	5.0	2.0	1.0
12-H2’	13-H1’	5.0	2.0	1.0
G13				
13-H1’	13-H8	3.8	2.0	0.7
13-H2’	13-H8	5.0	2.0	1.0
13-H3’	13-H8	3.8	2.0	0.7
13-H4’	13-H8	5.0	2.0	1.0
13-H5’	13-H8	5.0	2.0	1.0
13-H5"	13-H8	3.8	2.0	0.7
13-H1’	14-H8	5.0	3.0	1.0
13-H2’	14-H8	2.5	0.7	0.5
13-H3’	14-H8	3.8	2.0	0.7
13-H8	14-H8	5.0	2.0	1.0
13-H2’	14-H2’	5.0	2.0	1.0
13-H1’	14-H1’	5.0	2.0	1.0
13-H2’	14-H1’	5.0	2.0	1.0
G14				
14-H1’	14-H8	3.8	2.0	0.7
14-H2’	14-H8	5.0	2.0	1.0
14-H3’	14-H8	3.8	2.0	0.7
14-H4’	14-H8	5.0	2.0	1.0
14-H5’	14-H8	5.0	2.0	1.0
14-H5"	14-H8	3.8	2.0	0.7
14-H1’	15-H1’	5.0	2.0	1.0
14-H2’	15-H1’	5.0	2.0	1.0
14-H2’	15-H2’	5.0	2.0	1.0
14-H1’	15-H5	6.0	3.0	1.0
14-H2’	15-H5	3.8	2.0	0.7
14-H3’	15-H5	3.8	2.0	0.7
14-H8	15-H5	5.0	3.0	1.0
14-H1’	15-H5"	6.0	2.0	1.0
14-H1’	15-H6	5.0	2.0	1.0
14-H2’	15-H6	2.5	0.7	0.5
14-H3’	15-H6	3.8	2.0	0.7
14-H8	15-H6	5.0	2.0	1.0
U15				
15-H1’	15-H5	6.0	3.0	1.0
15-H2’	15-H5	5.0	2.0	1.0
15-H3’	15-H5	5.0	3.0	1.0
15-H1’	15-H6	3.8	2.0	0.7
15-H2’	15-H6	3.8	2.0	0.7
15-H3’	15-H6	3.8	2.0	0.7
15-H4’	15-H6	5.0	2.0	1.0
15-H5’	15-H6	5.0	3.0	1.0
15-H5"	15-H6	5.0	3.0	1.0
15-H1’	16-H1’	6.0	3.0	1.0
15-H2’	16-H1’	3.8	2.0	0.7
15-H3’	16-H1’	6.0	3.0	1.0
15-H2’	16-H2’	5.0	2.0	1.0
15-H1’	16-H8	5.0	2.0	1.0
15-H2’	16-H8	2.5	0.7	0.5
15-H3’	16-H8	3.8	2.0	0.7
15-H4’	16-H8	6.0	3.0	1.0
15-H5	16-H8	5.0	2.0	1.0
15-H6	16-H8	5.0	2.0	1.0
15-H1’	17-H8	6.0	2.0	2.0
A16				
16-H1’	16-H2	5.0	2.0	1.0
16-H1’	16-H8	3.8	2.0	0.7
16-H2’	16-H8	3.8	2.0	0.7
16-H3’	16-H8	3.8	2.0	0.7
16-H4’	16-H8	5.0	2.0	1.0
16-H5’	16-H8	5.0	3.0	1.0
16-H5"	16-H8	5.0	3.0	1.0
16-H1’	17-H1’	6.0	2.0	3.0
16-H2	17-H1’	5.0	3.0	1.0
16-H2’	17-H1’	6.0	2.0	1.0
16-H2	17-H2’	6.0	2.0	1.0
16-H1’	17-H3’	6.0	3.0	2.0
16-H1’	17-H8	6.0	3.0	1.0
16-H2’	17-H8	5.0	3.0	1.0
16-H3’	17-H8	5.0	3.0	1.0
16-H8	17-H8	5.0	3.0	1.0
16-H2	18-H1’	5.0	2.0	1.0
16-H2	18-H2’	6.0	2.0	2.0
16-H2’	18-H5	6.0	2.0	2.0
16-H2	18-H5	6.0	2.0	2.0
16-H2	19-H1’	6.0	3.0	1.0
16-H2	19-H5	3.8	2.0	0.7
16-H1’	20-H1’	6.0	3.0	1.0
G17				
17-H1’	17-H8	3.8	2.0	0.7
17-H2’	17-H8	3.8	2.0	0.7
17-H3’	17-H8	5.0	3.0	1.0
17-H1’	18-H1’	6.0	2.0	1.0
17-H1’	18-H2’	6.0	2.0	2.0
17-H1’	18-H3’	6.0	3.0	1.0
17-H8	18-H2’	6.0	2.0	1.0
17-H1’	18-H5	3.8	2.0	0.7
17-H2’	18-H5	6.0	2.0	1.0
17-H3’	18-H5	6.0	3.0	1.0
17-H8	18-H5	5.0	3.0	1.0
17-H1’	18-H6	5.0	3.0	1.0
17-H2’	18-H6	5.0	2.0	1.0
17-H3’	18-H6	6.0	3.0	1.0
17-H1’	19-H1’	6.0	2.0	1.0
U18				
18-H1’	18-H5	5.0	2.0	1.0
18-H2’	18-H5	3.8	2.0	0.7
18-H3’	18-H5	5.0	2.0	1.0
18-H4’	18-H5	6.0	2.0	2.0
18-H1’	18-H6	3.8	2.0	0.7
18-H2’	18-H6	2.5	0.7	0.5
18-H3’	18-H6	5.0	3.0	1.0
18-H4’	18-H6	5.0	3.0	1.0
18-H5’	18-H6	5.0	3.0	1.0
18-H5"	18-H6	5.0	3.0	1.0
18-H1’	19-H1’	6.0	2.0	3.0
18-H3’	19-H1’	6.0	2.0	3.0
18-H1’	19-H2’	6.0	2.0	2.0
18-H1’	19-H5	3.8	2.0	0.7
18-H2’	19-H5	5.0	3.0	1.0
18-H3’	19-H5	6.0	2.0	3.0
18-H4’	19-H5	6.0	3.0	1.0
18-H1’	19-H6	5.0	2.0	1.0
18-H1’	20-H1’	6.0	3.0	2.0
U19				
19-H1’	19-H5	5.0	2.0	1.0
19-H2’	19-H5	5.0	3.0	1.0
19-H1’	19-H6	3.8	2.0	0.7
19-H2’	19-H6	2.5	0.7	0.5
19-H4’	19-H6	5.0	3.0	1.0
19-H5’	19-H6	5.0	3.0	1.0
19-H5"	19-H6	5.0	3.0	1.0
19-H1’	20-H1’	6.0	3.0	1.0
19-H2’	20-H1’	5.0	3.0	1.0
19-H6	20-H1’	5.0	2.0	1.0
19-H1’	20-H5	6.0	3.0	1.0
19-H2’	20-H5	5.0	3.0	1.0
19-H1’	20-H6	3.8	2.0	0.7
19-H2’	20-H6	2.5	0.7	0.5
19-H6	20-H6	5.0	2.0	1.0
C20				
20-H1’	20-H6	3.8	2.0	0.7
20-H2’	20-H6	3.8	2.0	0.7
20-H3’	20-H6	3.8	2.0	0.7
20-H5"	20-H6	3.8	2.0	0.7
20-H1’	20-H5	5.0	2.0	1.0
20-H2’	20-H5	5.0	2.0	1.0
20-H3’	20-H5	5.0	2.0	1.0
20-H1’	21-H6	5.0	2.0	1.0
20-H2’	21-H6	2.5	0.7	0.5
20-H3’	21-H6	3.8	2.0	0.7

20-H6	21-H6	5.0	2.0	1.0
20-H2’	21-H5	3.8	2.0	0.7
20-H3’	21-H5	3.8	2.0	0.7
20-H5	21-H5	5.0	3.0	1.0
20-H2’	21-H2’	5.0	2.0	1.0
20-H1’	21-H1’	5.0	2.0	1.0
20-H2’	21-H1’	5.0	2.0	1.0
C21				
21-H1’	21-H6	3.8	2.0	0.7
21-H2’	21-H6	3.8	2.0	0.7
21-H3’	21-H6	3.8	2.0	0.7
21-H5"	21-H6	3.8	2.0	0.7
21-H1’	21-H5	5.0	2.0	1.0
21-H3’	21-H5	5.0	2.0	1.0
21-H1’	22-H8	4.5	2.0	0.7
21-H2’	22-H8	2.5	0.7	0.5
21-H3’	22-H8	3.8	2.0	0.7
21-H5	22-H8	6.0	2.0	1.0
21-H6	22-H8	5.0	2.0	1.0
21-H1’	22-H1’	5.0	2.0	1.0
21-H2’	22-H1’	5.0	2.0	1.0
G22				
22-H1’	22-H8	3.8	2.0	0.7
22-H2’	22-H8	5.0	2.0	1.0
22-H3’	22-H8	3.8	2.0	0.7
22-H4’	22-H8	5.0	2.0	1.0
22-H5’	22-H8	3.8	2.0	0.7
22-H5"	22-H8	3.8	2.0	0.7
22-H1’	23-H6	5.0	2.0	1.0
22-H2’	23-H6	2.5	0.7	0.5
22-H3’	23-H6	3.8	2.0	0.7
22-H1’	23-H5	5.0	2.0	1.0
22-H2’	23-H5	3.8	2.0	0.7
22-H3’	23-H5	5.0	2.0	1.0
22-H8	23-H5	5.0	2.0	1.0
22-H2’	23-H2’	5.0	2.0	1.0
22-H1’	23-H1’	5.0	2.0	1.0
22-H2’	23-H1’	5.0	2.0	1.0
C23				
23-H1’	23-H6	5.0	2.0	1.0
23-H2’	23-H6	5.0	2.0	1.0
23-H3’	23-H6	3.8	2.0	0.7
23-H5"	23-H6	5.0	2.0	1.0
23-H1’	23-H5	5.0	2.0	1.0
23-H2’	23-H5	5.0	2.0	1.0
23-H3’	23-H5	5.0	2.0	1.0
23-H5"	23-H5	6.0	2.0	1.0
23-H1’	24-H8	3.8	2.0	0.7
23-H2’	24-H8	2.5	0.7	0.5
23-H3’	24-H8	3.8	2.0	0.7
23-H6	24-H8	5.0	2.0	1.0
23-H2’	24-H2’	5.0	2.0	1.0
23-H1’	24-H1’	6.0	2.0	1.0
23-H2’	24-H1’	5.0	2.0	1.0
A24				
24-H1’	24-H8	3.8	2.0	0.7
24-H2’	24-H8	3.8	2.0	0.7
24-H3’	24-H8	3.8	2.0	0.7
24-H4’	24-H8	5.0	2.0	1.0
24-H5’	24-H8	5.0	2.0	1.0
24-H5"	24-H8	5.0	2.0	1.0
24-H1’	24-H2	5.0	2.0	1.0
24-H2’	24-H2	6.0	2.0	2.0
24-H1’	25-H6	5.0	2.0	1.0
24-H2	25-H6	6.0	2.0	2.0
24-H2’	25-H6	3.8	2.0	0.7
24-H3’	25-H6	3.8	2.0	0.7
24-H8	25-H6	5.0	2.0	1.0
24-H2’	25-H5	5.0	3.0	1.0
24-H8	25-H5	5.0	2.0	1.0
24-H1’	25-H2’	6.0	2.0	2.0
24-H2’	25-H2’	5.0	3.0	1.0
24-H1’	25-H1’	6.0	2.0	1.0
24-H2	25-H1’	3.8	2.0	0.7
24-H2’	25-H1’	6.0	2.0	1.0
C25				
25-H1’	25-H6	3.8	2.0	0.7
25-H2’	25-H6	2.5	0.7	0.5
25-H3’	25-H6	3.8	2.0	0.7
25-H4’	25-H6	5.0	2.0	1.0
25-H5’	25-H6	5.0	2.0	1.0
25-H5"	25-H6	5.0	2.0	1.0
25-H2’	25-H5	5.0	3.0	1.0
25-H1’	26-H8	5.0	2.0	1.0
25-H2’	26-H8	5.0	2.0	1.0
25-H4’	26-H8	5.0	2.0	1.0
25-H6	26-H8	6.0	2.0	1.0
25-H2’	26-H1’	6.0	1.0	3.0
25-H2’	26-H2’	6.0	1.0	3.0
G26				
26-H1’	26-H8	3.0	1.20	0.7
26-H2’	26-H8	5.0	2.0	1.0
26-H3’	26-H8	5.0	3.0	1.0
26-H4’	26-H8	5.0	2.0	1.0
26-H5’	26-H8	5.0	3.0	1.0
26-H5"	26-H8	3.8	2.0	0.7
26-H2’	27-H6	3.8	2.0	0.7
26-H2’	27-H5	5.0	3.0	1.0
26-H3’	27-H5	5.0	2.0	1.0
26-H2’	27-H2’	6.0	2.0	1.0
26-H2’	27-H1’	5.0	2.0	1.0
26-H8	28-H1’	6.0	2.0	1.0
U27				
27-H1’	27-H6	3.8	2.0	0.7
27-H2’	27-H6	3.8	2.0	0.7
27-H3’	27-H6	3.8	2.0	0.7
27-H5"	27-H6	5.0	2.0	1.0
27-H2’	27-H5	5.0	2.0	1.0
27-H3’	27-H5	5.0	2.0	1.0
27-H1’	28-H8	5.0	2.0	1.0
27-H2’	28-H8	2.5	0.7	0.5
27-H3’	28-H8	3.8	2.0	0.7
27-H6	28-H8	5.0	2.0	1.0
27-H2’	28-H2’	6.0	2.0	1.0
27-H1’	28-H1’	6.0	2.0	1.0
27-H2’	28-H1’	5.0	2.0	1.0
A28				
28-H1’	28-H8	3.8	2.0	0.7
28-H2’	28-H8	5.0	2.0	1.0
28-H3’	28-H8	3.8	2.0	0.7
28-H4’	28-H8	5.0	2.0	1.0
28-H5’	28-H8	5.0	2.0	1.0
28-H5"	28-H8	5.0	2.0	1.0
28-H1’	28-H2	5.0	2.0	1.0
28-H2’	28-H2	6.0	2.0	1.0
28-H1’	29-H6	5.0	2.0	1.0
28-H2	29-H6	6.0	2.0	1.0
28-H2’	29-H6	2.5	0.7	0.5
28-H3’	29-H6	3.8	2.0	0.7
28-H8	29-H6	5.0	2.0	1.0
28-H1’	29-H5	5.0	2.0	1.0
28-H2’	29-H5	3.8	2.0	0.7
28-H3’	29-H5	5.0	2.0	1.0
28-H8	29-H5	5.0	2.0	1.0
28-H1’	29-H4’	6.0	2.0	1.0
28-H2	29-H2’	6.0	2.0	1.0
28-H2’	29-H2’	5.0	2.0	1.0
28-H1’	29-H1’	5.0	2.0	1.0
28-H2	29-H1’	3.8	2.0	0.7
28-H2’	29-H1’	5.0	3.0	1.0
28-H3’	29-H1’	6.0	2.0	2.0
C29				
29-H1’	29-H6	3.8	2.0	0.7
29-H2’	29-H6	3.8	2.0	0.7
29-H3’	29-H6	3.8	2.0	0.7
29-H4’	29-H6	5.0	2.0	1.0
29-H5’	29-H6	3.8	2.0	0.7
29-H5"	29-H6	3.8	2.0	0.7
29-H1’	29-H5	5.0	2.0	1.0
29-H2’	29-H5	5.0	2.0	1.0
29-H3’	29-H5	5.0	2.0	1.0
29-H5"	29-H5	6.0	2.0	1.0
29-H1’	30-H8	5.0	2.0	1.0
29-H2’	30-H8	2.5	0.7	0.5
29-H3’	30-H8	3.8	2.0	0.7
29-H2’	30-H2’	5.0	2.0	1.0
29-H1’	30-H1’	5.0	2.0	1.0
29-H2’	30-H1’	5.0	3.0	1.0
G30				
30-H1’	30-H8	3.8	2.0	0.7
30-H2’	30-H8	5.0	2.0	1.0
30-H3’	30-H8	3.8	2.0	0.7
30-H4’	30-H8	5.0	2.0	1.0
30-H5’	30-H8	3.8	2.0	0.7
30-H5"	30-H8	3.8	2.0	0.7
30-H1’	31-H8	5.0	3.0	1.0
30-H2’	31-H8	2.5	0.7	0.5
30-H3’	31-H8	5.0	2.0	1.0
30-H8	31-H8	5.0	2.0	1.0
30-H2’	31-H1’	5.0	2.0	1.0
C34				
34-H1’	34-H6	3.8	2.0	0.7
34-H2’	34-H6	5.0	2.0	1.0
34-H3’	34-H6	3.8	2.0	0.7

34-H5"	34-H6	5.0	3.0	1.0
34-H2’	34-H5	5.0	3.0	1.0
34-H3’	34-H5	5.0	3.0	1.0
34-H1’	35-H6	5.0	2.0	1.0
34-H2’	35-H6	2.5	0.7	0.5
34-H3’	35-H6	3.8	2.0	0.7
34-H6	35-H6	5.0	2.0	1.0
34-H1’	35-H5	5.0	2.0	1.0
34-H2’	35-H5	5.0	2.0	1.0
34-H3’	35-H5	5.0	2.0	1.0
34-H5	35-H5	5.0	2.0	1.0
34-H6	35-H5	5.0	2.0	1.0
34-H2’	35-H2’	5.0	2.0	1.0
34-H1’	35-H1’	5.0	2.0	1.0
34-H2’	35-H1’	5.0	2.0	1.0
U35				
35-H1’	35-H6	3.8	2.0	0.7
35-H2’	35-H6	5.0	2.0	1.0
35-H3’	35-H6	3.8	2.0	0.7
35-H4’	35-H6	5.0	2.0	1.0
35-H5’	35-H6	3.8	2.0	0.7
35-H5"	35-H6	3.8	2.0	0.7
35-H1’	35-H5	5.0	2.0	1.0
35-H2’	35-H5	5.0	2.0	1.0
35-H3’	35-H5	5.0	2.0	1.0

Exchangeable protons:

G1				
1-H1	2-H1’	5.0	3.0	1.0
1-H1	34-H2’	6.0	3.0	1.0
1-H1	35-H2’	5.0	3.0	1.0
1-H1	35-H3	5.0	3.0	1.0
1-H1	35-H5	5.0	3.0	1.0
1-H22	35-H3	5.0	3.0	1.0
C6				
6-H5	7-H1	6.0	3.0	1.0
6-H41	29-H41	5.0	3.0	1.0
6-H1’	30-H1	6.0	3.0	1.0
6-H2’	30-H1	6.0	3.0	1.0
6-H42	30-H1	3.8	2.0	0.7
6-H41	30-H1	2.5	0.7	0.5
6-H5	30-H1	5.0	3.0	1.0
6-H6	30-H1	6.0	3.0	1.0
6-H1’	31-H1	6.0	3.0	1.0
6-H42	31-H1	5.0	3.0	1.0
6-H41	31-H1	5.0	3.0	1.0
6-H5	31-H1	5.0	3.0	1.0
6-H6	31-H1	6.0	3.0	1.0
G7				
7-H1’	7-H1	6.0	3.0	1.0
7-H1	8-H1’	6.0	3.0	1.0
7-H1	8-H3	5.0	3.0	1.0
7-H22	8-H3	6.0	3.0	1.0
7-H1	28-H2	5.0	3.0	1.0
7-H1	28-H61	5.0	3.0	1.0
7-H1	28-H62	6.0	3.0	1.0
7-H1	29-H1’	5.0	3.0	1.0
7-H1	29-H2’	6.0	2.0	1.0
7-H1	29-H42	3.8	2.0	0.7
7-H1	29-H41	2.5	0.7	0.5
7-H22	29-H41	6.0	2.0	1.0
7-H1	29-H5	5.0	3.0	1.0
7-H1	29-H6	6.0	3.0	1.0
7-H1	30-H1	5.0	3.0	1.0
7-H1	30-H1’	6.0	2.0	1.0
7-H1’	30-H1	5.0	3.0	1.0
U8				
8-H3	9-H8	6.0	3.0	1.0
8-H3	28-H2	3.8	2.0	0.7
8-H3	28-H62	3.8	2.0	0.7
8-H3	28-H61	2.5	0.7	0.5
8-H3	29-H42	6.0	3.0	1.0
8-H3	29-H41	5.0	3.0	1.0
G11				
11-H1	12-H1’	6.0	3.0	1.0
11-H1	12-H42	5.0	3.0	1.0
11-H1	12-H41	5.0	3.0	1.0
11-H1	12-H5	6.0	2.0	1.0
11-H1	22-H1	3.8	2.0	0.7
11-H1	23-H1’	5.0	3.0	1.0
11-H1	23-H2’	6.0	2.0	1.0
11-H1	23-H42	3.8	2.0	0.7
11-H1	23-H41	2.5	0.7	0.5
11-H1	23-H5	5.0	3.0	1.0
11-H1	23-H6	6.0	3.0	1.0
11-H1	24-H1’	5.0	3.0	1.0
11-H1	24-H2	5.0	3.0	1.0
C12				
12-H41	13-H1	5.0	3.0	1.0
12-H42	13-H1	5.0	3.0	1.0
12-H1’	22-H1	5.0	3.0	1.0
12-H2’	22-H1	6.0	3.0	1.0
12-H42	22-H1	3.8	2.0	0.7
12-H41	22-H1	2.5	0.7	0.5
12-H5	22-H1	5.0	2.0	1.0
12-H6	22-H1	6.0	2.0	1.0
12-H41	23-H41	6.0	3.0	1.0
G13				
13-H1’	13-H1	6.0	2.0	1.0
13-H1	14-H1	3.8	2.0	0.7
13-H1	14-H1’	6.0	3.0	1.0
13-H1	20-H2’	6.0	3.0	1.0
13-H1	20-H42	6.0	3.0	1.0
13-H1	21-H1’	6.0	3.0	1.0
13-H1	21-H2’	6.0	3.0	1.0
13-H1	21-H42	3.8	2.0	0.7
13-H1	21-H41	2.5	0.7	0.5
13-H1	21-H5	5.0	3.0	1.0
13-H1	21-H6	6.0	3.0	1.0
13-H1	22-H1	5.0	3.0	1.0
13-H1	22-H1’	6.0	3.0	1.0
13-H1	22-H8	6.0	3.0	1.0
13-H8	22-H1	6.0	3.0	1.0
G14				
14-H1’	14-H1	6.0	3.0	1.0
14-H2’	14-H1	6.0	3.0	1.0
14-H1	15-H1’	5.0	3.0	1.0
14-H1	20-H1’	6.0	3.0	1.0
14-H1	20-H2’	6.0	2.0	1.0
14-H1	20-H42	3.8	2.0	0.7
14-H1	20-H41	2.5	0.7	0.5
14-H1	20-H5	5.0	3.0	1.0
14-H1	20-H6	6.0	3.0	1.0
14-H1	21-H1’	5.0	3.0	1.0
14-H1	21-H42	5.0	2.0	1.0
14-H1	21-H5	5.0	3.0	1.0
14-H1	21-H6	6.0	3.0	1.0
14-H21	15-H1’	6.0	3.0	1.0
14-H21	21-H1’	5.0	3.0	1.0
U15				
14-H1	15-H3	5.0	3.0	1.0
15-H3	15-H1’	6.0	3.0	1.0
15-H3	15-H5	5.0	3.0	1.0
15-H3	15-H6	6.0	3.0	1.0
15-H3	16-H1’	6.0	3.0	1.0
15-H3	16-H8	6.0	3.0	1.0
15-H3	19-H5	6.0	3.0	1.0
15-H3	20-H41	5.0	3.0	1.0
15-H3	20-H42	5.0	3.0	1.0
C20				
20-H42	21-H42	5.0	3.0	1.0
20-H42	21-H5	6.0	3.0	1.0
20-H6	21-H42	6.0	3.0	1.0
C21				
21-H2’	21-H41	6.0	2.0	1.0
21-H41	22-H1	5.0	2.0	1.0
21-H42	22-H1	6.0	2.0	1.0
21-H5	22-H1	6.0	2.0	1.0
G22				
22-H1’	22-H1	6.0	2.0	1.0
22-H2’	22-H1	6.0	2.0	1.0
22-H1	23-H1’	6.0	3.0	1.0
22-H1	23-H42	5.0	3.0	1.0
22-H1	23-H41	5.0	3.0	1.0
C29				
29-H41	30-H1	5.0	3.0	1.0
G30				
30-H1’	30-H1	6.0	3.0	1.0
30-H1	31-H1	5.0	3.0	1.0
30-H1	31-H1’	6.0	3.0	1.0
C34				
34-H42	35-H3	6.0	3.0	1.0
34-H41	35-H3	5.0	3.0	1.0
34-H5	35-H3	6.0	3.0	1.0

NOE-derived distance restraints in the catalytic triad, pH 5.2-5.5:

Non-exchangeable protons:

A3				
3-H1’	3-H8	5.0	3.0	1.0
3-H1’	3-H2	6.0	2.0	1.0
3-H2	4-H1’	5.0	3.0	1.0
3-H2	4-H8	5.0	3.0	1.0
3-H2	32-H2	5.0	3.0	1.0
3-H61	32-H2	5.0	3.0	1.0
3-H2	34-H1’	5.0	3.0	1.0
3-H2	34-H2’	6.0	3.0	1.0
3-H61	34-H41	5.0	3.0	1.0
3-H2	34-H6	6.0	3.0	1.0
G4				
4-H1’	4-H8	3.8	2.0	0.7
4-H1’	5-H6	5.0	2.0	1.0
4-H2’	5-H6	5.0	3.0	1.0
4-H1’	5-H5	5.0	2.0	1.0
4-H8	5-H1’	5.0	3.0	1.0
4-H1’	5-H1’	6.0	3.0	1.0
4-H2’	5-H1’	5.0	3.0	1.0
4-H1’	32-H2	6.0	3.0	1.0
4-H8	32-H2	3.8	2.0	0.7
C5				
5-H1’	5-H6	5.0	3.0	1.0
5-H2’	5-H6	5.0	3.0	1.0
5-H3’	5-H6	5.0	3.0	1.0
5-H1’	5-H5	6.0	3.0	1.0
5-H2’	5-H5	5.0	3.0	1.0
5-H1’	6-H6	5.0	3.0	1.0
5-H2’	6-H6	2.5	0.7	0.5
5-H3’	6-H6	3.8	2.0	0.7
5-H2’	6-H5	3.8	2.0	0.7
5-H5	6-H5	3.8	2.0	0.7
5-H6	6-H5	6.0	3.0	1.0
5-H2’	6-H2’	6.0	2.0	1.0
5-H1’	6-H1’	5.0	2.0	1.0
5-H2’	6-H1’	5.0	3.0	1.0
5-H1’	32-H2	3.8	2.0	0.7
G31				
31-H1’	31-H8	3.8	2.0	0.7
31-H2’	31-H8	5.0	3.0	1.0
31-H3’	31-H8	3.8	2.0	0.7
31-H4’	31-H8	5.0	2.0	1.0
31-H5’	31-H8	5.0	2.0	1.0
31-H5"	31-H8	3.8	2.0	0.7
31-H1’	32-H8	5.0	3.0	1.0
31-H2’	32-H8	2.5	0.7	0.5
31-H3’	32-H8	5.0	3.0	1.0
31-H8	32-H8	6.0	3.0	1.0
31-H1’	32-H1’	6.0	3.0	1.0
31-H2’	32-H1’	5.0	3.0	1.0
A32				
32-H1’	32-H8	3.8	2.0	0.7
32-H2’	32-H8	5.0	3.0	1.0
32-H5"	32-H8	5.0	3.0	1.0
32-H1’	32-H2	5.0	2.0	1.0
32-H1’	33-H6	5.0	3.0	1.0
32-H2’	33-H6	2.5	0.7	0.5
32-H3’	33-H6	3.8	2.0	0.7
32-H8	33-H6	6.0	3.0	1.0
32-H1’	33-H5	5.0	3.0	1.0
32-H2’	33-H5	5.0	3.0	1.0
32-H3’	33-H5	5.0	3.0	1.0
32-H2	33-H5	6.0	3.0	1.0
32-H8	33-H5	6.0	3.0	1.0
32-H1’	33-H1’	5.0	2.0	1.0
32-H2	33-H1’	3.8	2.0	0.7
32-H2’	33-H1’	5.0	3.0	1.0
U33				
33-H1’	33-H6	5.0	3.0	1.0
33-H2’	33-H6	5.0	3.0	1.0
33-H3’	33-H6	5.0	3.0	1.0
33-H5"	33-H6	5.0	3.0	1.0
33-H1’	33-H5	6.0	3.0	1.0
33-H3’	33-H5	5.0	3.0	1.0
33-H1’	34-H6	5.0	3.0	1.0
33-H2’	34-H6	3.8	2.0	0.7
33-H3’	34-H6	3.8	2.0	0.7
33-H6	34-H6	5.0	2.0	1.0
33-H3’	34-H5	5.0	2.0	1.0
33-H5	34-H5	5.0	3.0	1.0

Exchangeable protons:

G2				
2-H1	3-H2	5.0	3.0	1.0
2-H1	3-H61	5.0	3.0	1.0
2-H1	33-H3	5.0	3.0	1.0
2-H1	34-H42	3.8	2.0	0.7
2-H1	34-H41	2.5	0.7	0.5
2-H1	34-H5	5.0	3.0	1.0
2-H1	35-H3	5.0	3.0	1.0
2-H1	35-H6	5.0	3.0	1.0
A3				
3-H8	4-H1	5.0	3.0	1.0
3-H61	32-H61	5.0	3.0	2.0
3-H61	32-H62	5.0	3.0	2.0
3-H62	32-H61	5.0	3.0	2.0
3-H62	32-H62	5.0	3.0	2.0
3-H61	33-H3	5.0	3.0	1.0
3-H62	33-H3	5.0	3.0	1.0
G4				
4-H1	5-H42	6.0	3.0	1.0
4-H1’	33-H3	6.0	3.0	1.0
C5				
5-H41	6-H41	5.0	3.0	1.0
5-H41	6-H5	5.0	3.0	1.0
5-H41	30-H1	5.0	3.0	1.0
5-H42	31-H1	5.0	3.0	1.0
5-H41	31-H1	3.8	2.0	1.0
5-H5	31-H1	5.0	3.0	1.0
5-H5	32-H61	5.0	3.0	1.0
5-H5	32-H62	6.0	3.0	1.0
G31				
31-H1	32-H61	5.0	3.0	1.0
31-H1	32-H62	5.0	3.0	1.0
A32				
32-H2	33-H3	5.0	3.0	1.0
32-H61	33-H3	6.0	3.0	1.0
32-H61	33-H3	6.0	3.0	1.0
U33				
33-H1’	33-H3	6.0	3.0	1.0
33-H3	34-H41	5.0	3.0	1.0
33-H3	34-H42	5.0	3.0	1.0

NOE-derived distance restraints in the catalytic triad, pH 7.8-7.9:

Non-exchangeable protons:

A3

3-H2	4-H1’	6.0	3.0	1.0
3-H2	34-H1’	5.0	3.0	1.0
3-H2	34-H6	6.0	3.0	1.0

G4

4-H1’	4-H8	5.0	3.0	1.0
4-H1’	5-H6	6.0	3.0	1.0
4-H2’	5-H6	5.0	3.0	1.0
4-H1’	5-H5	5.0	2.0	1.0
4-H1’	5-H1’	6.0	3.0	1.0

C5

5-H1’	5-H6	5.0	3.0	1.0
5-H2’	5-H6	5.0	3.0	1.0
5-H3’	5-H6	5.0	3.0	1.0
5-H1’	5-H5	6.0	3.0	1.0
5-H2’	5-H5	5.0	3.0	1.0
5-H1’	6-H6	5.0	3.0	1.0
5-H2’	6-H6	2.5	0.7	0.5
5-H3’	6-H6	3.8	2.0	0.7
5-H2’	6-H5	5.0	3.0	1.0
5-H5	6-H5	5.0	3.0	1.0
5-H6	6-H5	6.0	3.0	1.0
5-H2’	6-H2’	6.0	2.0	1.0
5-H1’	6-H1’	5.0	2.0	1.0
5-H2’	6-H1’	5.0	3.0	1.0
5-H1’	32-H2	3.8	2.0	0.7

G31

31-H1’	31-H8	3.8	2.0	0.7
31-H2’	31-H8	5.0	3.0	1.0
31-H3’	31-H8	5.0	3.0	1.0
31-H5"	31-H8	5.0	3.0	1.0
31-H1’	32-H8	5.0	3.0	1.0
31-H2’	32-H8	5.0	3.0	1.0
31-H3’	32-H8	5.0	3.0	1.0
31-H8	32-H8	6.0	3.0	1.0
31-H1’	32-H1’	6.0	3.0	1.0

A32

32-H1’	32-H8	3.8	2.0	0.7
32-H2’	32-H8	5.0	3.0	1.0
32-H5"	32-H8	5.0	3.0	1.0
32-H1’	32-H2	5.0	2.0	1.0
32-H1’	33-H6	5.0	3.0	1.0
32-H2’	33-H6	5.0	3.0	1.0
32-H3’	33-H6	5.0	3.0	1.0
32-H2’	33-H5	5.0	3.0	1.0
32-H3’	33-H5	5.0	3.0	1.0
32-H1’	33-H1’	6.0	2.0	1.0
32-H2	33-H1’	3.8	2.0	0.7
32-H2’	33-H1’	5.0	3.0	1.0

U33

33-H1’	33-H6	5.0	3.0	1.0
33-H2’	33-H6	5.0	3.0	1.0
33-H3’	33-H6	5.0	3.0	1.0
33-H5"	33-H6	5.0	3.0	1.0
33-H3’	33-H5	5.0	3.0	1.0
33-H1’	34-H6	5.0	3.0	1.0
33-H2’	34-H6	3.8	2.0	0.7
33-H3’	34-H6	3.8	2.0	0.7
33-H6	34-H6	5.0	2.0	1.0
33-H3’	34-H5	5.0	2.0	1.0
33-H5	34-H5	5.0	3.0	1.0

Exchangeable protons:

G2

2-H1	3-H2	5.0	3.0	1.0
2-H1	34-H42	3.8	2.0	0.7
2-H1	34-H41	2.5	0.7	0.5
2-H1	34-H5	5.0	3.0	1.0
2-H1	35-H3	5.0	3.0	1.0
2-H1	35-H6	5.0	3.0	1.0

C5

5-H41	6-H41	5.0	3.0	1.0
5-H41	6-H5	5.0	3.0	1.0
5-H41	30-H1	5.0	3.0	1.0
5-H42	31-H1	5.0	3.0	1.0
5-H41	31-H1	3.8	2.0	1.0
5-H5	31-H1	5.0	3.0	1.0

Appendix 11. Dihedral angles for standard A-form RNA [13,428]. In structure calculations of AvD5 and D1kz angles were restrained to these values in the helical regions. For details on dihedral restraints in non-helical regions see Experimental Section 4.5.

backbone	
α (O3'$_{i-1}$-P$_i$-O5'$_i$-C5'$_i$)	-68°
β (P$_i$-O5'$_i$-C5'$_i$-C4'$_i$)	178°
γ (O5'$_i$-C5'$_i$-C4'$_i$-C3'$_i$)	54°
ϵ (C4'$_i$-C3'$_i$-O3'$_i$-P$_{i+1}$)	-153°
ζ (C3'$_i$-O3'$_i$-P$_{i+1}$-O5'$_{i+1}$)	-71°

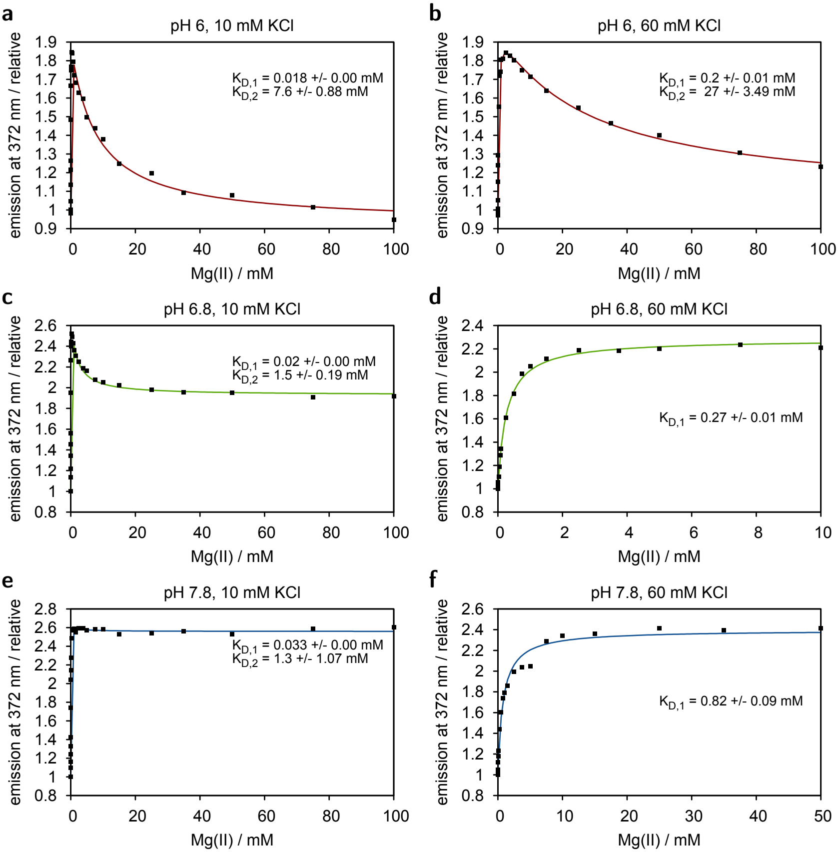
sugar pucker	(C2'-endo / C3'-endo)
δ (C5'$_i$-C4'$_i$-C3'$_i$-O3'$_i$)	145° / 85°
ν_1 (O4'$_i$-C1'$_i$-C2'$_i$-C3'$_i$)	25° / -25°
ν_2 (C1'$_i$-C2'$_i$-C3'$_i$-C4'$_i$)	-35° / 37°

glycosidic angle	(anti / syn)
χ (O4'$_i$-C1'$_i$-N9/N1$_i$-C4/C2$_i$)	-160° / 60°

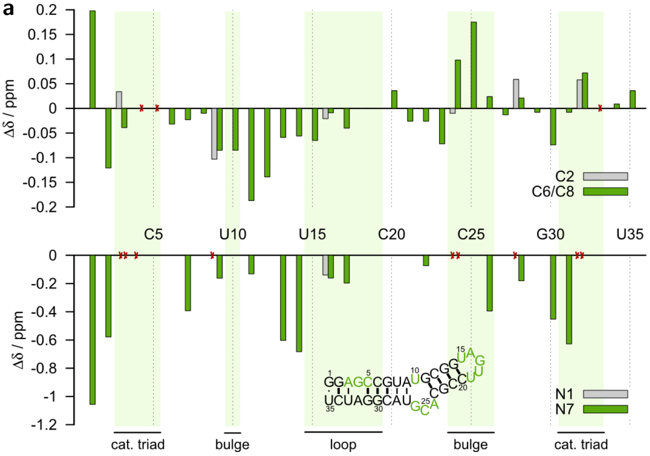
Appendix 12. RDC restraints used in AvD5 structure calculations.

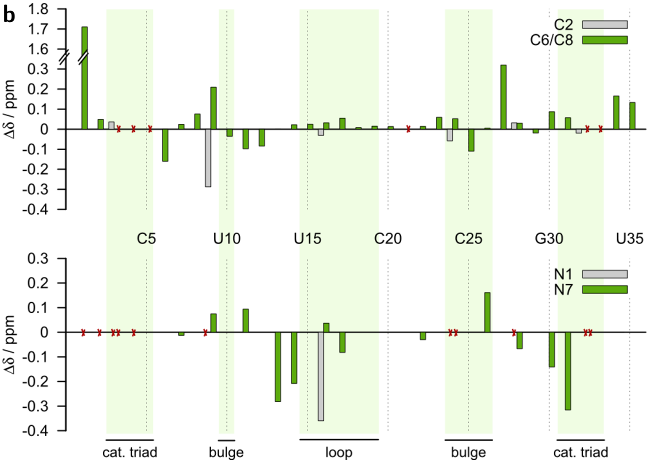
first nucleus	second nucleus	RDC (Hz)	RDC error (Hz)
8-C1’	8-H1’	-7	1.5
10-C1’	10-H1’	-13	1.5
16-C1’	16-H1’	-8	1.5
20-C1’	20-H1’	-22	1.5
21-C1’	21-H1’	-20	1.5
16-C2	16-H2	11	1.5
28-C2	28-H2	18	1.5
10-C5	10-H5	16	1.5
34-C5	34-H5	23	1.5
10-C6	10-H6	27	1.5
20-C6	20-H6	22	1.5
34-C6	34-H6	18	1.5
2-C8	2-H8	23	1.5
11-C8	11-H8	25	1.5
13-C8	13-H8	21	1.5
14-C8	14-H8	20	1.5
16-C8	16-H8	2	1.5
28-C8	28-H8	22	1.5
30-C8	30-H8	21	1.5
31-C8	31-H8	25	1.5
7-N1	7-H1	31	3.0
8-N3	8-H3	31	3.0
11-N1	11-H1	31	3.0
13-N1	13-H1	22	3.0
14-N1	14-H1	22	3.0
22-N1	22-H1	30	3.0

Appendix 13. Mono- and biphasic curve fits of the fluorescence decrease in AvD5-AP24 upon addition of Mg(II) at 10 and 60 mM KCl and pH 6, 6.8 and 7.8. Measurements were done in 10 mM MOPS, 1 µM RNA at room temperature. Equations for single- and two-site binding were applied as appropriate for the shape of the curves (Equations 4.3 and 4.4). When there is only a very slight decrease in emission at high Mg(II) concentrations the data is not fit well by the biphasic model and was fit with the monophasic equation. In such cases, when necessary, the last points of the titration were not included in the fitting (Figure d,f).

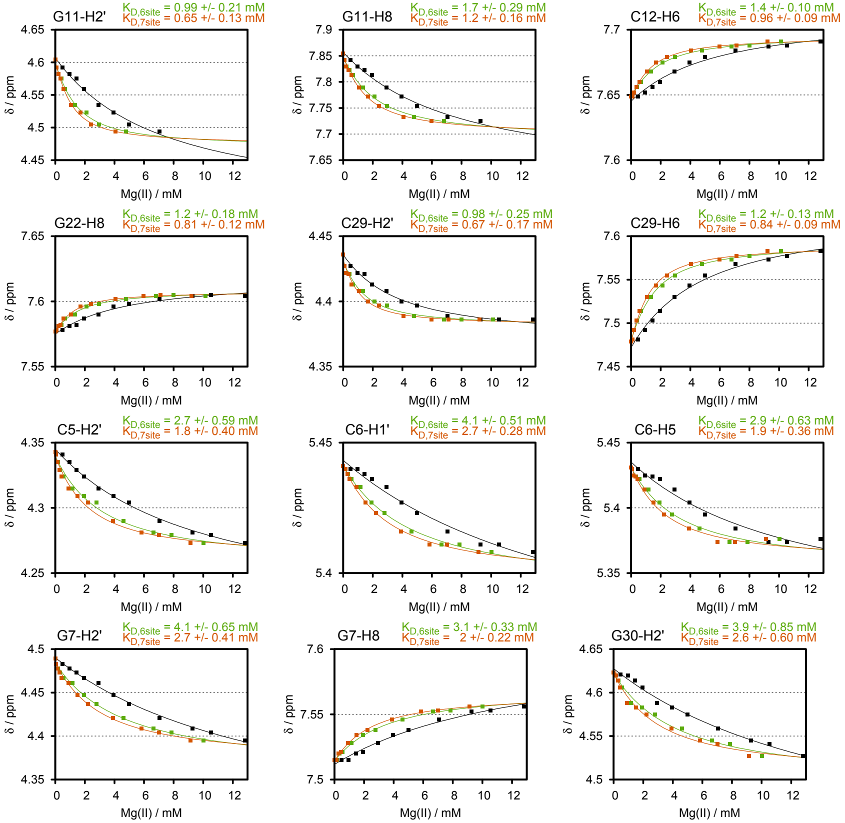


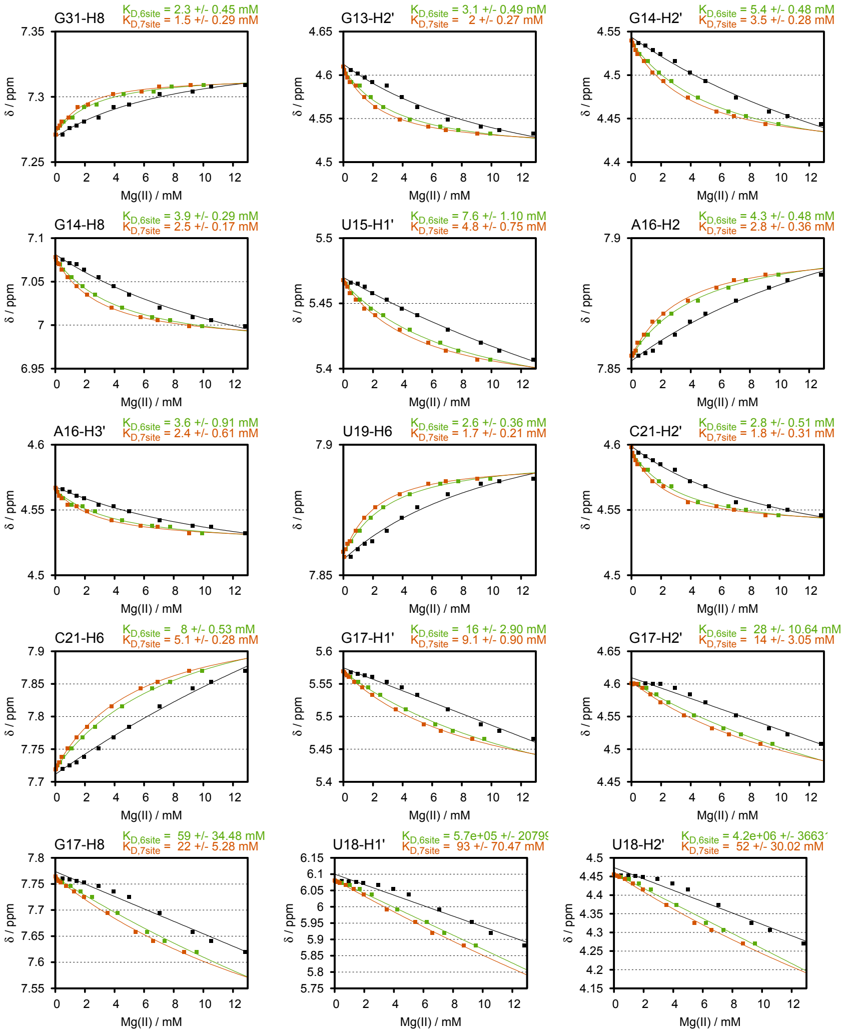
Appendix 14. AvD5 ^{13}C and ^{15}N chemical shift perturbations in the presence of Mg(II) and Cd(II). a) ^{13}C and ^{15}N chemical shift changes upon addition of 2.5 mM of $MgCl_2$ (0.8 mM AvD5, 290 K, pD 6.7, 60 mM KCl, 10 µM EDTA). b) ^{13}C and ^{15}N chemical shift changes in AvD5 upon addition of 0.5 mM of $Cd(ClO_4)_2$ (0.8 mM AvD5, 290 K, pD 6.7, 60 mM $KClO_4$, 10 µM EDTA). To distinguish resonances that are not moving at all from values that are missing because assignment is not possible under all conditions the latter were crossed out red in the graph. Catalytic triad, bulge and pentaloop are green in the secondary structure scheme.

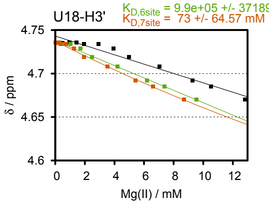


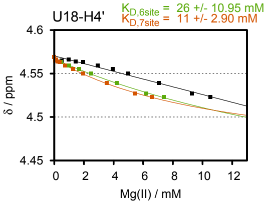


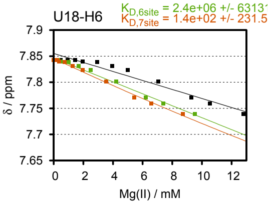
Appendix 15. Least-squares fits (Equation 4.3 in the Experimental Section) of Mg(II)-dependent ^{1}H chemical shift changes in AvD5. In the black curves the chemical shift (δ) is plotted against the added amount of Mg(II), in the green and orange curves they are plotted against the actually freely available amount of Mg(II) in the presence of a total of 6 or 7 binding sites, respectively. For details see Section 2.5.1. K_Ds derived from the fits for the green and orange curves are indicated at the top. They are also found in Tables 2.5 and 2.6 ($K_{D,fin,6sites}$ and $K_{D,fin,7sites}$, respectively), together with the K_D values from fitting the black curves ($K_{D,ini}$).

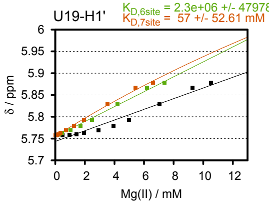


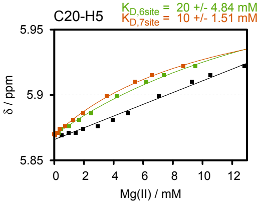

G31-H8
$K_{D,6site}$ = 2.3 +/- 0.45 mM
$K_{D,7site}$ = 1.5 +/- 0.29 mM
7.35
7.3
7.25
G13-H2'
$K_{D,6site}$ = 3.1 +/- 0.49 mM
$K_{D,7site}$ = 2 +/- 0.27 mM
4.65
4.6
4.55
4.5
G14-H2'
$K_{D,6site}$ = 5.4 +/- 0.48 mM
$K_{D,7site}$ = 3.5 +/- 0.28 mM
4.55
4.5
4.45
4.4
G14-H8
$K_{D,6site}$ = 3.9 +/- 0.29 mM
$K_{D,7site}$ = 2.5 +/- 0.17 mM
7.1
7.05
7
6.95
U15-H1'
$K_{D,6site}$ = 7.6 +/- 1.10 mM
$K_{D,7site}$ = 4.8 +/- 0.75 mM
5.5
5.45
5.4
A16-H2
$K_{D,6site}$ = 4.3 +/- 0.48 mM
$K_{D,7site}$ = 2.8 +/- 0.36 mM
7.9
7.85
A16-H3'
$K_{D,6site}$ = 3.6 +/- 0.91 mM
$K_{D,7site}$ = 2.4 +/- 0.61 mM
4.6
4.55
4.5
U19-H6
$K_{D,6site}$ = 2.6 +/- 0.36 mM
$K_{D,7site}$ = 1.7 +/- 0.21 mM
7.9
7.85
C21-H2'
$K_{D,6site}$ = 2.8 +/- 0.51 mM
$K_{D,7site}$ = 1.8 +/- 0.31 mM
4.6
4.55
4.5
C21-H6
$K_{D,6site}$ = 8 +/- 0.53 mM
$K_{D,7site}$ = 5.1 +/- 0.28 mM
7.9
7.85
7.8
7.75
7.7
G17-H1'
$K_{D,6site}$ = 16 +/- 2.90 mM
$K_{D,7site}$ = 9.1 +/- 0.90 mM
5.6
5.55
5.5
5.45
5.4
G17-H2'
$K_{D,6site}$ = 28 +/- 10.64 mM
$K_{D,7site}$ = 14 +/- 3.05 mM
4.65
4.6
4.55
4.5
4.45
G17-H8
$K_{D,6site}$ = 59 +/- 34.48 mM
$K_{D,7site}$ = 22 +/- 5.28 mM
7.8
7.75
7.7
7.65
7.6
7.55
U18-H1'
$K_{D,6site}$ = 5.7e+05 +/- 2079
$K_{D,7site}$ = 93 +/- 70.47 mM
6.15
6.1
6.05
6
5.95
5.9
5.85
5.8
5.75
U18-H2'
$K_{D,6site}$ = 4.2e+06 +/- 3663
$K_{D,7site}$ = 52 +/- 30.02 mM
4.5
4.45
4.4
4.35
4.3
4.25
4.2
4.15
δ / ppm
0
2
4
6
8
10
12
Mg(II) / mM

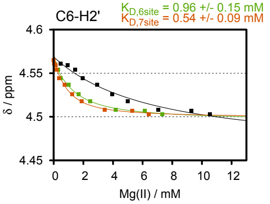

U18-H3'
$K_{D,6site}$ = 9.9e+05 +/- 37189
$K_{D,7site}$ = 73 +/- 64.57 mM
4.75
4.7
4.65
4.6
δ / ppm
0 2 4 6 8 10 12
Mg(II) / mM

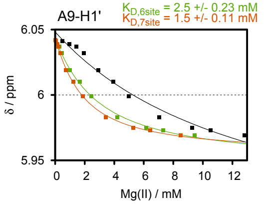

U18-H4'
$K_{D,6site}$ = 26 +/- 10.95 mM
$K_{D,7site}$ = 11 +/- 2.90 mM
4.6
4.55
4.5
4.45
δ / ppm
0 2 4 6 8 10 12
Mg(II) / mM

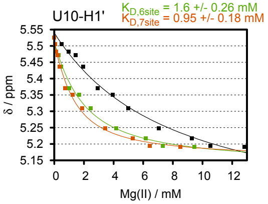

U18-H6
$K_{D,6site}$ = 2.4e+06 +/- 63131
$K_{D,7site}$ = 1.4e+02 +/- 231.5
7.9
7.85
7.8
7.75
7.7
7.65
δ / ppm
0 2 4 6 8 10 12
Mg(II) / mM

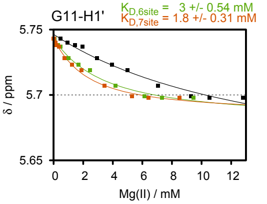

U19-H1'
$K_{D,6site}$ = 2.3e+06 +/- 47978
$K_{D,7site}$ = 57 +/- 52.61 mM
6
5.95
5.9
5.85
5.8
5.75
5.7
δ / ppm
0 2 4 6 8 10 12
Mg(II) / mM

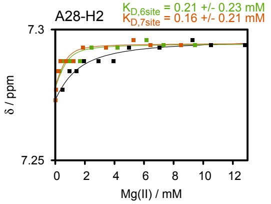

C20-H5
$K_{D,6site}$ = 20 +/- 4.84 mM
$K_{D,7site}$ = 10 +/- 1.51 mM
5.95
5.9
5.85
δ / ppm
0 2 4 6 8 10 12
Mg(II) / mM

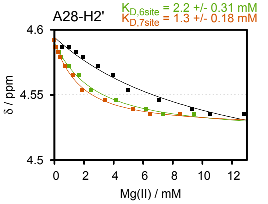

C6-H2'
$K_{D,6site}$ = 0.96 +/- 0.15 mM
$K_{D,7site}$ = 0.54 +/- 0.09 mM
4.6
4.55
4.5
4.45
δ / ppm
0 2 4 6 8 10 12
Mg(II) / mM

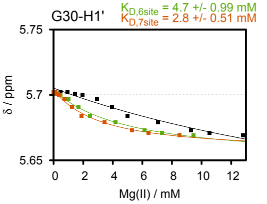

A9-H1'
$K_{D,6site}$ = 2.5 +/- 0.23 mM
$K_{D,7site}$ = 1.5 +/- 0.11 mM
6.05
6
5.95
δ / ppm
0 2 4 6 8 10 12
Mg(II) / mM

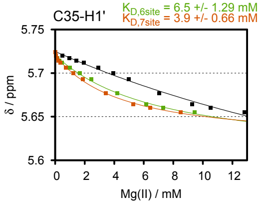

U10-H1'
$K_{D,6site}$ = 1.6 +/- 0.26 mM
$K_{D,7site}$ = 0.95 +/- 0.18 mM
5.55
5.5
5.45
5.4
5.35
5.3
5.25
5.2
5.15
δ / ppm
0 2 4 6 8 10 12
Mg(II) / mM


G11-H1'
$K_{D,6site}$ = 3 +/- 0.54 mM
$K_{D,7site}$ = 1.8 +/- 0.31 mM
5.75
5.7
5.65
δ / ppm
0 2 4 6 8 10 12
Mg(II) / mM


A28-H2
$K_{D,6site}$ = 0.21 +/- 0.23 mM
$K_{D,7site}$ = 0.16 +/- 0.21 mM
7.3
7.25
δ / ppm
0 2 4 6 8 10 12
Mg(II) / mM


A28-H2'
$K_{D,6site}$ = 2.2 +/- 0.31 mM
$K_{D,7site}$ = 1.3 +/- 0.18 mM
4.6
4.55
4.5
δ / ppm
0 2 4 6 8 10 12
Mg(II) / mM


G30-H1'
$K_{D,6site}$ = 4.7 +/- 0.99 mM
$K_{D,7site}$ = 2.8 +/- 0.51 mM
5.75
5.7
5.65
δ / ppm
0 2 4 6 8 10 12
Mg(II) / mM


C35-H1'
$K_{D,6site}$ = 6.5 +/- 1.29 mM
$K_{D,7site}$ = 3.9 +/- 0.66 mM
5.75
5.7
5.65
5.6
δ / ppm
0 2 4 6 8 10 12
Mg(II) / mM

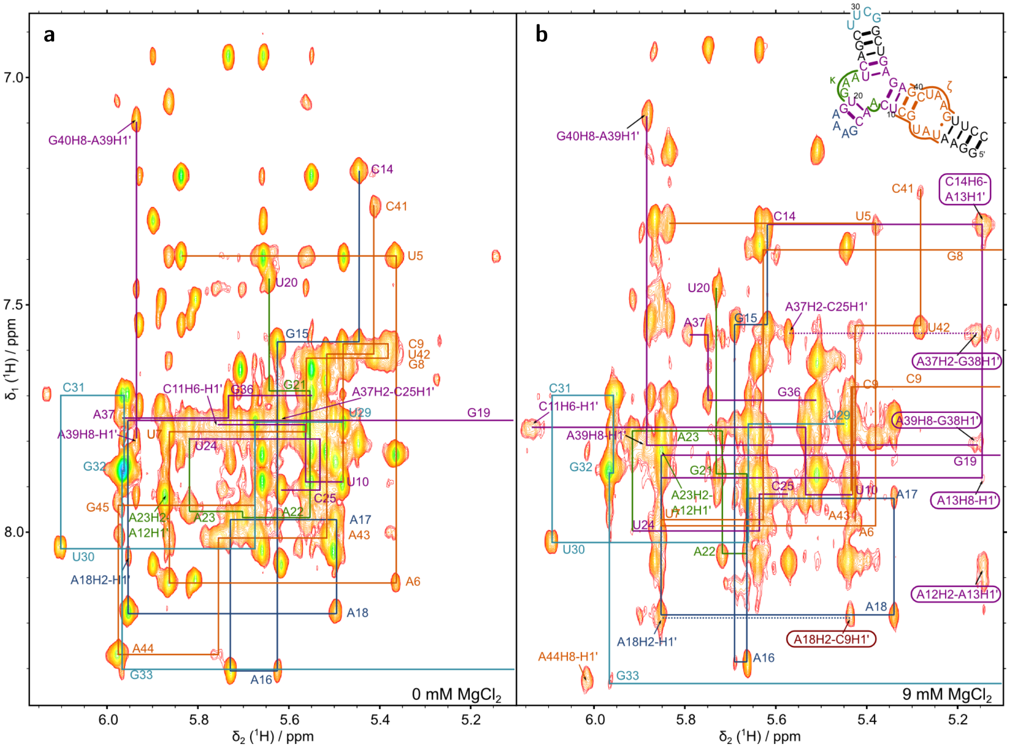
Appendix 16. D1-27 ^{1}H chemical shifts. Non-exchangeable proton resonances were recorded at 300 K, 3 mM D1-27, 900 MHz, exchangeable ones at 275 K, 0.75 mM D1-27, 600 MHz, all in 60 mM KCl, 10 µM EDTA, pH/pD 6.75. ^{a}Ade-H2/Pyr-H5, ^{b}Pyr-H6/Pur-H8, ^{c}Gua-H21/Cyt-H41/Ade-H61, ^{d}Gua-H22/Cyt-H42/Ade-H62, ^{e}Ura-H3/Gua-H1.

	H1'	H2'	H3'	H4'	H5'	H5"	H2/H5^{a}	H6/H8^{b}	amino-H1^{c}	amino-H2^{d}	imino-H^{e}
G1	5.837	4.947	4.713	4.573	4.446	4.291	-	8.139	-	-	12.728
G2	5.907	4.703	4.620	4.544	4.512	4.274	-	7.548	8.099	5.914	12.376
A3	5.954	4.688	4.665	4.544	4.588	4.169	7.235	7.743	7.833	6.744	-
A4	5.861	4.520	4.408	4.501	4.534	4.107	7.849	7.560	8.065	6.781	-
U5	5.407	4.279	4.515	4.385	4.470	4.075	5.208	7.420	-	-	11.562
A6	5.903	4.632	4.796	4.566	4.366	4.186	7.440	8.135	-	-	-
U7	5.600	4.558	4.538	4.492	4.422	4.252	5.663	7.818	-	-	-
G8	5.404	4.516	4.456	4.495	4.401	4.164	-	7.640	8.581	6.312	13.146
C9	5.530	4.417	4.535	4.454	4.498	4.112	5.301	7.648	8.722	6.991	-
U10	5.589	4.594	4.470	4.530	4.576	4.107	5.437	7.943	-	-	14.083
C11	5.557	4.440	4.308	4.467	4.522	4.108	5.657	7.781	8.467	7.030	-
U12	5.533	3.751	4.551	4.366	4.513	4.100	5.745	7.786	-	-	11.743
U13	6.102	4.659	4.023	4.489	4.249	4.045	5.869	8.031	-	-	-
C14	5.955	4.103	4.494	3.805	2.719	3.616	6.140	7.686	6.397	7.115	-
G15	5.973	4.854	5.643	4.423	4.201	4.403	-	7.868	6.663	-	9.820
G16	4.459	4.547	4.357	4.426	4.545	4.308	-	8.305	8.167	6.367	12.526
A17	5.981	4.697	4.678	4.459	4.545	4.095	7.493	7.776	8.047	6.544	-
G18	5.646	4.443	4.368	4.458	4.455	4.042	-	7.151	8.637	6.233	13.439
C19	5.520	4.438	4.265	4.384	4.473	4.047	5.149	7.387	8.366	7.024	-
U20	5.561	4.133	4.438	4.365	4.491	4.081	5.517	7.650	-	-	-
A21	5.788	4.341	4.693	4.326	4.269	4.109	7.079	8.038	-	-	-
A22	6.000	4.883	4.930	4.619	4.429	4.205	8.025	8.286	-	-	-
G23	5.681	4.712	4.561	4.569	4.434	4.265	-	7.944	6.378	-	11.239
U24	5.519	4.433	4.497	4.500	4.586	4.101	5.373	7.758	-	-	14.083
U25	5.680	4.482	4.561	4.525	4.580	4.138	5.600	8.054	-	-	13.806
C26	5.570	4.211	4.489	4.436	4.577	4.099	5.671	7.907	8.355	7.037	-
C27	5.726	3.995	4.178	4.166	4.464	4.039	5.495	7.658	8.283	7.014	-

Appendix 17. D1-27 ^{13}C chemical shifts. RNA concentration 3 mM, 60 mM KCl, 10 µM EDTA, pD 6.75, 900 MHz. ^{a}Ade-C2/Pyr-C5, ^{b}Pyr-C6/Pur-C8.

	C1'	C2'	C3'	C4'	C5'	C2/C5^{a}	C6/C8^{b}
G1	91.764	75.407	75.153	83.509	67.598	-	139.088
G2	92.711	75.745	73.438	82.789	66.377	-	137.068
A3	93.209	76.126	73.079	82.604	65.362	153.336	139.887
A4	93.246	76.024	72.931	82.355	65.526	154.159	139.167
U5	92.97	75.907	73.46	83.16	65.166	104.328	140.594
A6	90.129	76.799	76.9	84.699	67.342	154.799	141.473
U7	92.582	75.764	73.128	84.308	67.13	104.839	143.136
G8	92.331	75.502	73.249	83.424	65.772	-	137.137
C9	94.237	75.692	74.035	82.336	65.493	97.424	141.565
U10	93.921	75.46	72.629	82.591	64.757	103.662	142.365
C11	94.094	75.631	72.165	82.092	64.79	98.068	141.785
U12	94.555	76.149	73.406	82.958	64.822	105.169	141.154
U13	89.414	74.811	78.004	87.07	67.949	105.664	144.836
C14	89.292	77.755	80.542	84.649	67.661	98.778	143.305
G15	94.684	77.461	76.172	83.393	69.152	-	143.074
G16	93.297	75.257	74.85	83.322	69.971	-	139.117
A17	93.061	75.787	72.673	82.379	64.831	153.464	139.87
G18	92.785	75.631	73.382	82.404	65.879	-	135.805
C19	94.562	76.022	72.64	82.532	65.091	97.533	140.974
U20	93.873	75.816	73.458	83.278	65.313	104.75	141.224
A21	90.092	77.122	76.879	85.237	67.124	154.213	141.202
A22	90.969	76.294	76.238	84.866	67.314	154.832	142.006
G23	92.342	75.554	74.075	83.664	66.997	-	138.586
U24	94.082	75.506	72.719	82.612	65.154	103.686	141.707
U25	94.204	75.549	72.638	84.244	64.682	103.619	143.063
C26	94.352	75.786	72.474	82.414	64.805	97.542	142.121
C27	92.916	77.804	70.035	83.764	65.502	98.258	141.745

Appendix 18. D1kz sequential walk in the absence (a) and presence (b) of 9 mM $MgCl_2$.



Connections of the helices d and d” are not labelled in the spectra to reduce crowding. The sections of the sequential walk have been colour-coded according to the secondary structure scheme on top. Newly appearing resonances are marked by boxes in (b). They are situated mainly in the nucleotides directly neighbouring the three-way junction, the only exception being A18H2-C9H1’. (0.6-0.7 mM D1kz, 300 K, 60 mM KCl, 10 µM EDTA, pD 6.7-6.8, 700 MHz).

Appendix 19. D1kz ^{1}H chemical shifts in the absence of Mg(II). Non-exchangeable protons were recorded at 300 K, exchangeable ones at 285 K. RNA concentrations 0.4-1 mM, pH 6.6-6.8, 60 mM KCl, 10 µM EDTA, 600 and 700 MHz. ^{a}Ade-H2/Pyr-H5, ^{b}Pyr-H6/Pur-H8, ^{c}Gua-H21/Cyt-H41/Ade-H61, ^{d}Gua-H22/Cyt-H42/Ade-H62, ^{e}Ura-H3/Gua-H1.

	H1'	H2'	H3'	H4'	H5'	H5"	H2/H5^{a}	H6/H8^{b}	amino-H1^{c}	amino-H2^{d}	imino-H^{e}
G1	5.811	4.923	4.701	4.512	4.307	4.153	-	8.104	-	-	12.774
G2	5.882	4.678	4.607	4.532	4.503	4.229	-	7.49	-	-	12.368
A3	5.929	4.665	4.635	-	4.608	4.143	7.22	7.724	-	-	-
A4	5.839	4.505	4.38	4.483	-	4.081	7.828	7.543	-	-	-
U5	5.364	4.238	4.475	4.362	-	4.082	5.19	7.394	-	-	11.569
A6	5.864	4.581	4.766	4.543	4.369	4.141	7.401	8.112	-	-	-
U7	5.563	4.514	-	-	-	4.217	5.617	7.78	-	-	-
G8	5.382	4.496	4.42	-	-	4.083	-	7.622	-	-	13.123
C9	5.483	4.419	4.369	-	-	4.084	5.266	7.587	8.627	6.912	-
U10	5.562	4.435	4.479	-	-	-	5.383	7.89	-	-	14.326
C11	5.76	4.32	-	-	-	-	5.646	7.772	-	-	-
A12	5.872	4.565	-	-	-	-	7.967	7.819	-	-	-
A13	5.409	4.408	4.502	-	-	-	7.808	7.932	-	-	-
C14	5.451	4.295	4.28	4.329	4.401	4.044	5.079	7.209	8.096	6.752	-
G15	5.629	4.478	4.658	4.436	-	4.08	-	7.589	-	-	10.55
A16	5.727	4.727	4.514	4.235	3.979	4.176	7.849	8.307	-	-	-
A17	5.499	4.47	4.601	4.303	-	3.966	7.841	7.974	-	-	-
A18	5.952	4.646	5.002	4.426	4.6	4.296	8.057	8.174	-	-	-
G19	-	4.241	4.428	4.259	4.276	4.261	-	7.747	-	-	12.901
U20	5.645	4.375	4.321	-	-	-	5.09	7.442	-	-	-
G21	5.555	4.541	-	-	-	-	-	7.688	-	-	-
A22	5.703	4.552	-	-	-	-	7.764	7.974	-	-	-
A23	5.825	4.559	-	-	-	-	7.921	7.954	-	-	-
U24	5.54	4.489	-	-	-	4.215	5.571	7.792	-	-	13.344
C25	5.618	4.527	4.58	4.524	-	4.213	5.743	7.907	8.319	7.002	-
A26	5.897	4.621	4.749	4.525	-	4.196	6.954	8.068	-	-	-
G27	5.658	4.427	4.453	4.487	4.521	4.081	-	7.315	-	-	13.445
C28	5.481	4.461	4.229	4.401	-	4.038	5.143	7.397	8.569	6.994	-
U29	5.675	3.793	4.526	4.368	4.502	4.082	5.714	7.759	-	-	11.79
U30	6.106	4.677	4.026	4.484	4.24	4.028	5.861	8.031	-	-	-
C31	5.961	4.101	4.485	3.797	2.714	3.609	6.13	7.699	6.369	7.151	-
G32	5.966	4.836	5.634	4.409	4.192	4.414	-	7.864	6.677	-	9.919
G33	4.442	4.473	4.246	4.392	4.491	4.277	-	8.311	8.894	6.479	13.485

	H1’	H2’	H3’	H4’	H5’	H5"	H2/H5	H6/H8	amino-H1	amino-H2	imino-H
C34	5.52	4.464	4.417	-	-	4.02	5.276	7.664	8.741	6.934	-
U35	5.55	4.623	4.555	4.5	-	4.096	5.469	7.827	-	-	13.525
G36	5.733	4.532	4.515	4.5	-	4.129	-	7.697	-	-	11.748
A37	5.958	4.534	-	-	-	-	7.759	7.755	-	-	-
G38	5.312	4.619	-	-	-	-	-	7.482	-	-	-
A39	5.937	4.589	4.64	-	-	-	7.469	7.794	-	-	-
G40	-	4.295	4.307	-	-	4.004	-	7.093	-	-	13.384
C41	5.413	4.366	4.19	-	-	-	5.073	7.281	8.307	6.917	-
U42	5.517	4.082	4.407	4.32	-	-	5.473	7.616	-	-	-
A43	5.747	4.32	4.665	4.297	4.237	4.077	7.062	8.013	-	-	-
A44	5.983	4.88	4.921	4.599	4.402	4.179	8.003	8.267	-	-	-
G45	5.678	4.705	4.545	-	-	-	-	7.944	-	-	11.324
U46	5.505	4.416	4.48	4.472	-	-	5.357	7.75	-	-	14.048
U47	5.66	4.462	4.544	4.406	-	4.084	5.58	8.038	-	-	13.779
C48	5.552	4.191	4.472	-	-	4.079	5.654	7.89	8.371	6.998	-
C49	5.707	3.975	4.155	4.152	-	4.021	5.473	7.642	8.298	6.99	-

Appendix 20. D1kz ^{13}C and ^{15}N chemical shifts in the absence of Mg(II). All shifts were recorded at 300 K, only amino and imino ^{15}N shifts at 280 K. RNA concentrations 0.4-1 mM, pH 6.6-6.8, 60 mM KCl, 10 µM EDTA, 600 and 700 MHz. ^{a}Ade-C2/Pyr-C5, ^{b}Pyr-C6/Pur-C8.

	C1’	C2’	C3’	C4’	C5’	C2/C5^{a}	C6/C8^{b}	N1	N3	N7
G1	91.725	75.167	74.186	83.482	65.483	-	138.952	147.285	-	232.651
G2	92.785	75.401	72.89	82.205	65.637	-	136.716	146.922	-	234.25
A3	92.9	75.55	72.933	-	65.304	153.032	139.361	221.18	212.757	230.31
A4	92.892	75.317	72.795	-	-	153.945	138.938	222.507	213.716	231.502
U5	92.886	75.679	73.329	82.745	-	104.091	140.318	-	157.477	-
A6	90.106	75.327	-	84.749	-	154.603	141.158	224.575	215.853	232.608
U7	92.536	75.365	-	-	-	104.659	142.871	-	-	-
G8	92.133	75.402	-	-	-	-	136.979	148.412	-	233.575
C9	94.235	-	72.355	-	-	97.21	141.342	-	-	-
U10	94.017	-	73.203	-	-	103.153	141.977	-	-	-
C11	92.59	-	-	-	-	98.103	141.965	-	-	-
A12	91.46	-	-	-	-	155.029	-	223.987	216.531	-
A13	-	75.398	73.273	-	-	154.603	139.874	223.606	214.034	-

	C1’	C2’	C3’	C4’	C5’	C2/C5^a	C6/C8^b	N1	N3	N7
C14	93.734	75.727	72.624	82.031	-	97.512	140.19	-	-	-
G15	92.386	75.351	-	81.843	-	-	137.001	146.103	-	237.013
A16	91.22	75.974	74.559	83.896	65.538	154.52	142.121	224.78	216.212	230.254
A17	91.616	75.813	-	83.728	65.213	154.534	140.739	225.633	215.754	228.718
A18	91.744	76.509	74.655	82.929	-	155.207	141.246	226.038	217.276	230.994
G19	-	-	-	83.044	68.704	-	137.616	-	-	234.762
U20	92.53	75.78	74.901	-	-	103.718	141.223	-	-	-
G21	92.137	75.851	-	-	-	-	137.589	-	-	237.178
A22	92.397	-	-	-	-	154.617	140.848	223.906	215.788	230.194
A23	91.754	-	-	-	-	154.673	140.436	225.59	-	-
U24	93.913	-	-	-	-	104.309	-	-	-	-
C25	93.886	-	72.866	83.16	-	98.039	141.85	-	-	-
A26	92.639	75.355	73.031	-	-	152.505	139.614	220.31	213.086	230.613
G27	92.417	75.35	72.588	82.006	65.442	-	135.719	148.331	-	234.413
C28	93.859	75.42	71.902	81.636	64.674	97.257	140.325	-	-	-
U29	94.431	75.933	73.261	82.17	64.066	105.058	140.664	-	159.994	-
U30	89.189	74.583	77.731	86.957	67.749	105.484	144.684	-	-	-
C31	89.068	77.589	80.363	84.467	67.354	98.532	142.817	-	-	-
G32	94.54	77.32	75.894	83.155	69.009	-	142.945	143.28	-	231.955
G33	93.181	75.602	74.397	83.064	69.79	-	138.847	148.317	-	233.444
C34	93.999	75.197	71.96	-	-	96.925	141.322	-	-	-
U35	93.729	75.415	72.323	-	-	103.821	141.561	-	162.076	-
G36	92.632	75.202	-	-	-	-	136.349	-	-	235.449
A37	92.527	-	-	-	-	153.996	139.444	-	-	-
G38	90.097	-	-	-	-	-	-	-	-	-
A39	92.731	75.497	72.7	-	-	153.275	139.493	221.541	212.71	229.684
G40	-	76.938	-	-	-	-	135.53	148.127	-	234.46
C41	93.796	75.554	72.528	-	-	97.223	140.553	-	-	-
U42	93.782	75.562	73.425	82.742	-	104.389	141.432	-	-	-
A43	89.809	76.616	74.412	84.891	-	154.062	141.007	226.57	216.397	232.839
A44	90.734	75.994	-	84.69	66.861	154.688	141.882	225.613	216.522	232.869
G45	92.201	-	-	-	-	-	138.236	144.768	-	235.383
U46	93.855	75.611	72.518	82.346	-	103.426	141.617	-	163.081	-
U47	93.702	75.293	72.448	82.54	-	103.39	142.588	-	163.275	-
C48	94.107	75.656	72.215	-	-	97.462	141.881	-	-	-
C49	92.727	77.537	69.802	83.525	-	98.018	141.809	-	-	-

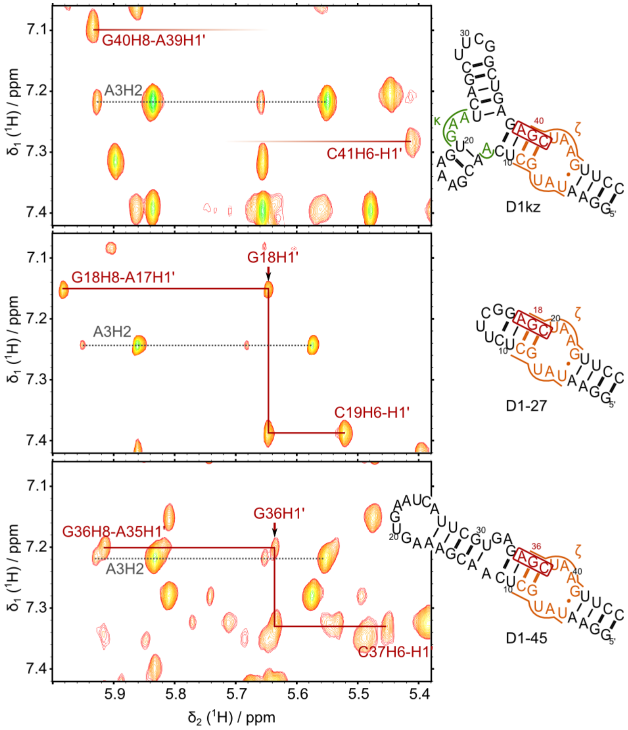
Appendix 21. D1kz ^{1}H chemical shifts in the presence of 10 mM MgCl$_2$. Non-exchangeable protons were recorded at 300 K, exchangeable ones at 285 K. RNA concentrations 0.4-1 mM, pH 6.6-6.8, 60 mM KCl, 10 µM EDTA, 600 and 700 MHz. [a]Pyr-H6/Pur-H8, [b]Gua-H21/Cyt-H41/Ade-H61, [c]Gua-H22/Cyt-H42/Ade-H62, [d]Ura-H3/Gua-H1.

	H1'	H2'	H2	H6/H8[a]	amino-H1[b]	amino-H2[c]	imino-H[d]
G1	5.685	4.812	-	8.116	-	-	12.713
G2	5.834	4.574	-	7.51	-	-	12.422
A3	5.875	4.526	7.16	7.705	7.885	6.487	-
A4	5.833	4.36	7.793	7.533	8.082	6.462	-
U5	5.379	4.373	-	7.323	-	-	11.234
A6	5.856	4.735	-	7.991	-	-	-
U7	5.637	4.619	-	7.987	-	-	-
G8	5.006	4.364	-	7.396	-	-	12.965
C9	5.432	4.316	-	7.684	8.517	6.845	-
U10	5.531	4.317	-	7.914	-	-	14.37
C11	6.138	4.282	-	7.757	8.395	6.835	-
A12	5.851	4.97	8.08	7.484	-	-	-
A13	5.143	4.231	7.924	7.871	-	-	-
C14	5.618	4.357	-	7.324	8.137	6.761	-
G15	5.689	4.271	-	7.545	-	8.221	10.772
A16	5.661	4.782	7.746	8.288	-	-	-
A17	5.34	4.537	7.77	7.919	-	-	-
A18	5.852	4.718	8.183	8.182	-	-	-
G19	3.689	3.98	-	7.84	-	-	13.069
U20	5.727	4.289	-	7.461	-	-	14.576
G21	5.662	4.477	-	7.87	-	-	11.396
A22	5.716	4.635	7.782	8.044	-	-	-
A23	5.914	4.514	7.818	7.778	-	-	-
U24	5.633	4.601	-	7.986	-	-	14.108
C25	5.57	4.479	-	7.917	8.181	6.942	-
A26	5.866	4.55	6.935	8.062	-	-	-
G27	5.634	4.382	-	7.31	-	6.289	13.476
C28	5.443	4.424	-	7.368	8.539	6.821	-
U29	5.659	3.774	-	7.762	-	-	11.802
U30	6.093	4.657	-	8.018	-	-	-
C31	5.955	4.093	-	7.695	6.383	7.14	-
G32	5.965	4.826	-	7.863	6.681	-	9.905
G33	4.42	4.434	-	8.331	8.88	6.525	13.478
C34	5.509	4.406	-	7.72	8.738	6.902	-
U35	5.514	4.576	-	7.869	-	-	13.513
G36	5.748	4.55	-	7.707	-	5.859	11.596
A37	5.792	4.667	7.562	7.563	-	-	-
G38	5.155	4.377	-	7.201	-	-	12.526
A39	5.882	4.476	7.571	7.807	-	-	-
G40	-	4.094	-	7.082	-	-	13.474
C41	5.282	4.334	-	7.253	8.312	6.729	-
U42	5.421	4.174	-	7.549	-	-	-
A43	5.772	4.364	6.982	7.978	-	-	-
A44	6.013	4.964	8.06	8.325	-	-	-
G45	5.508	4.564	-	7.82	6.384	-	11.173
U46	5.442	4.33	-	7.726	-	-	14.01
U47	5.64	4.416	-	8.063	-	-	13.776
C48	5.509	4.16	-	7.88	8.305	6.99	-
C49	5.639	3.953	-	7.625	8.361	6.987	-

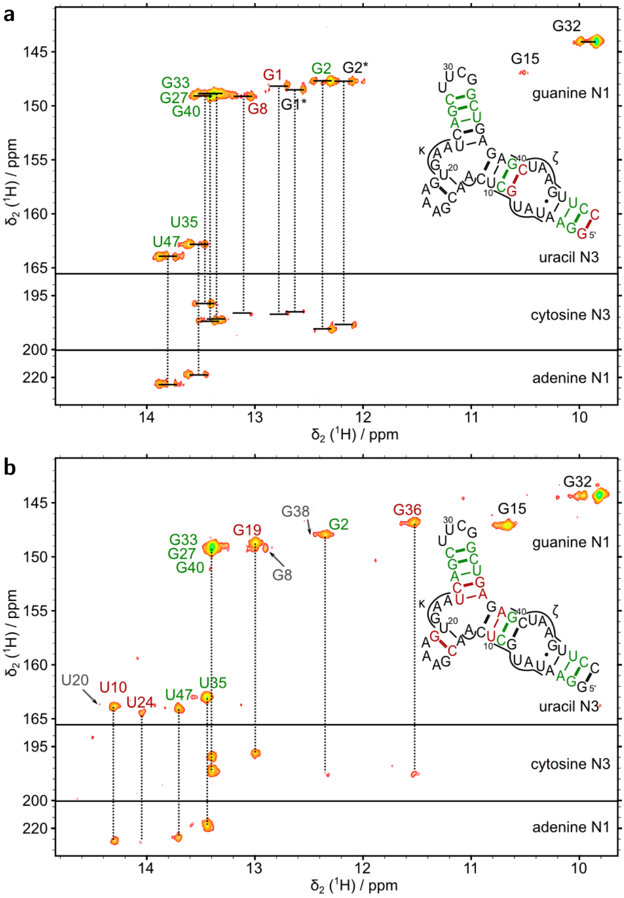
Appendix 22. D1kz ^{13}C and ^{15}N chemical shifts in the presence of 10 mM MgCl$_2$. All shifts were recorded at 300 K, only amino and imino ^{15}N shifts at 280 K. RNA concentrations 0.4-1 mM, pH 6.6-6.8, 60 mM KCl, 10 µM EDTA, 600 and 700 MHz. ^{a}Pyr-C6/Pur-C8.

	C1'	C2'	C2	C6/C8^a	N1	N3	N7
G1	-	-	-	-	148.32	-	-
G2	-	-	-	136.996	147.019	-	-
A3	-	-	153.116	139.276	221.203	-	-
A4	-	-	154.075	138.83	-	-	-
U5	-	-	-	140.291	-	158.331	-
A6	88.64	-	154.72	141.194	-	-	-
U7	-	-	-	143.267	-	-	-
G8	-	-	-	137.3	148.062	-	-
C9	-	-	-	141.095	-	-	-
U10	-	-	-	141.983	-	162.797	-
C11	91.907	-	-	-	-	-	-
A12	-	-	154.038	139.636	-	216.72	-
A13	92.357	-	154.005	139.153	-	-	-
C14	-	-	-	139.819	-	-	-
G15	-	-	-	136.6	146.081	-	-
A16	-	-	153.994	142.153	223.404	215.982	228.702
A17	93.059	-	154.42	140.067	226.233	215.181	-
A18	-	-	155.517	140.159	226.467	215.997	-
G19	93.082	-	-	137.297	147.726	-	-
U20	-	-	-	140.816	-	161.506	-
G21	-	-	-	137.87	-	-	-
A22	-	-	154.621	141.406	-	-	-
A23	-	-	-	140.26	-	-	-
U24	-	-	-	143.321	-	163.407	-
C25	-	-	-	141.868	-	-	-
A26	-	-	152.536	139.566	219.982	213.026	-
G27	-	-	-	135.676	148.245	-	-
C28	-	-	-	140.308	-	-	-
U29	-	75.9	-	140.711	-	160.062	-
U30	89.156	-	-	144.643	-	-	-
C31	89.078	77.527	-	142.865	-	-	-
G32	94.522	-	-	142.957	-	-	231.972
G33	93.206	-	-	138.93	148.435	-	-
C34	-	-	-	141.431	-	-	-
U35	-	-	-	141.466	-	161.908	-
G36	-	-	-	136.298	145.81	-	-
A37	-	-	153.562	139.127	-	-	-
G38	-	-	-	135.698	146.876	-	-
A39	-	-	153.532	139.298	-	-	-
G40	-	-	-	135.376	147.949	-	-
C41	93.659	-	-	140.457	-	-	-
U42	-	-	-	141.441	-	-	-
A43	-	-	153.655	141.108	-	-	-
A44	-	-	154.794	141.791	-	-	-
G45	-	-	-	-	144.655	-	-
U46	-	-	-	141.411	-	163.354	-
U47	-	-	-	142.69	-	162.878	-
C48	-	-	-	141.85	-	-	-
C49	-	-	-	141.517	-	-	-

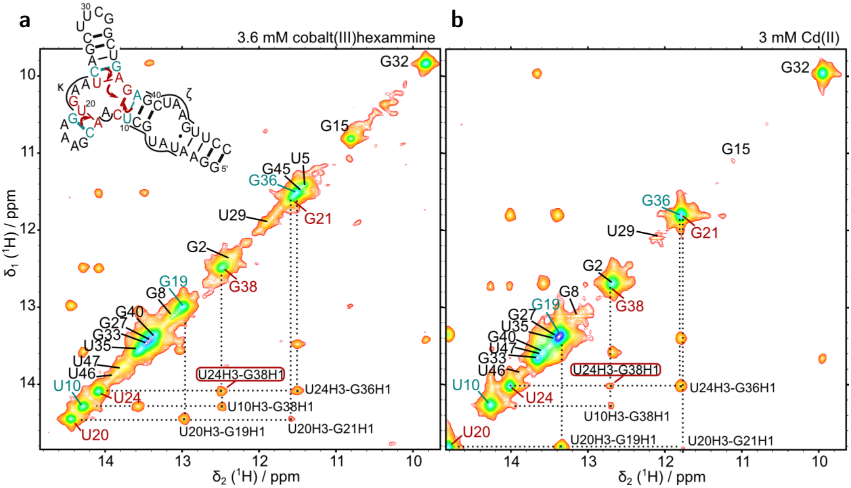
Appendix 23. Sections of ^{1}H,^{1}H-NOESY spectra of three different constructs containing ζ (orange), indicating that the absence of D1kz-G40H1' is caused by the adjacent κ element (green). Only in the absence of κ (green) the proton at the position equivalent to G40H1' in D1kz (G18H1' in D1-27 and G36H1' in D1-45) is observable (indicated in red in the spectra and secondary structure schemes). The relatively invariable A3H2 resonance is indicated in all spectra as a point of reference. (0.5-1 mM RNA, 300 K, 60 mM KCl, 10 µM EDTA, 700 MHz)



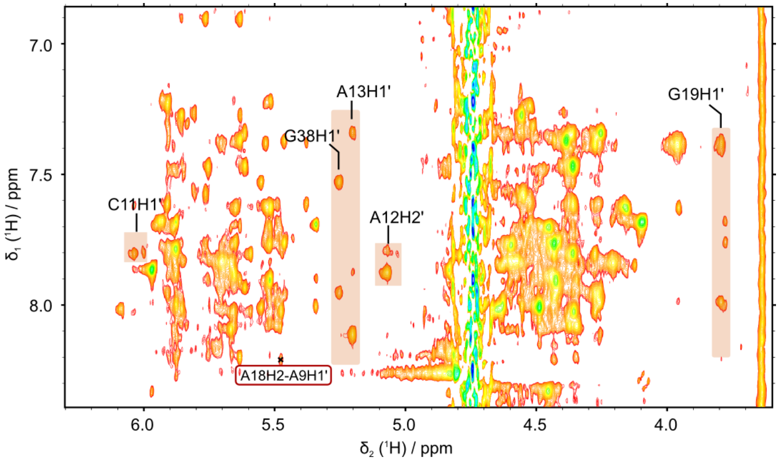
Appendix 24. J_{NN} HNN COSY spectra demonstrating the base pairing pattern in D1kz in the absence (a) and presence (b) of 12 mM $MgCl_2$. Base pairs that are observable under both conditions are coloured green, those that occur only under one of the two are red and base pairs without observable H-bonds are left black in the secondary structure schemes. Spectra were recorded using the TROSY version of the J_{NN} HNN COSY experiment[389]. In (a) the parameters were not optimized and therefore peaks are doubled. (conditions: a) 0.5 mM D1kz, 280 K, pH 6.9, 600 MHz; b) 1.1 mM D1kz, 285 K, pH 6.8, 700 MHz; 60 mM KCl, 10 µM EDTA)



Appendix 25. ^{1}H,^{1}H-NOESY imino proton region of D1kz in the presence of 3.6 mM cobalt(III)hexammine (a) and 3 mM Cd(II) (b). The same resonances and cross peaks appear as upon addition of Mg(II). Resonances that are not observable in the absence of Cd(II) or cobalt(III)hexammine are labelled in red, their neighbours are labelled in cyan. The crucial new cross peak from U24H3 to G38H1, which establishes a connection between two sides of the three-way junction, is boxed in red. Sequential NOE connectivities that appear in the presence of di- or trivalent metal ions are marked by red arrows in the secondary structure scheme. (conditions: a) 0.22 mM D1kz, 278 K, pH 6.7, 60 mM KCl, 10 µM EDTA, 700 MHz; b) 0.6 mM D1kz, 280 K, pH 6.7, 60 mM $KClO_4$, 10 µM EDTA, 700 MHz)



Appendix 26. $^1H,^1H$-NOESY of D1kz in the presence of 3.6 mM cobalt(III)hexammine indicating newly appearing non-exchangeable resonances at the κ three-way junction (C11H1’, A12H2’, A13H1’, G38H1’) and at the GAAA loop of the κ extension (G19H1’) (highlighted in orange). In addition, the intermolecular cross peak A18H2-A9H1’, indicating a D1kz dimer in the presence of cobalt(III)hexammine, is marked by a red box. (0.35 mM D1kz, 300 K, 60 mM KCl, 10 µM EDTA, pD 6.9, 700 MHz)



Appendix 27. NOE-derived distance restraints used in D1-27 structure calculations. Distance restraints derived from $^1H,^1H$-NOESY spectra recorded in D_2O at 300 K (non-exchangeable protons) and in H_2O at 275 K (exchangeable protons), pD/pH 6.7, 60 mM KCl, 10 µM EDTA, 600 and 700 MHz:

Restraints are given in the following format (distances in Å):

<first proton> <second proton> <target distance> <lower limit> <upper limit>

Non-exchangeable protons:

G1				
1-H8	1-H1’	3.8	2.0	0.7
1-H8	1-H2’	3.8	2.0	0.7
1-H8	1-H3’	3.8	2.0	0.7
1-H8	1-H4’	5.0	2.0	1.0
1-H8	1-H5”	3.8	2.0	0.7
1-H8	1-H5’	3.8	2.0	0.7
1-H1’	2-H8	3.8	2.0	0.7
1-H2’	2-H5”	5.0	2.0	1.0
1-H2’	2-H5’	3.8	2.0	0.7
1-H2’	2-H1’	5.0	2.0	1.0
1-H2’	2-H8	2.5	0.7	0.5
1-H3’	2-H8	3.8	2.0	0.7
1-H8	2-H8	5.0	2.0	1.0
G2				
2-H8	2-H1’	3.8	2.0	0.7
2-H8	2-H2’	5.0	2.0	1.0
2-H8	2-H3’	3.8	2.0	0.7
2-H8	2-H4’	5.0	2.0	1.0
2-H8	2-H5”	3.8	2.0	0.7
2-H8	2-H5’	3.8	2.0	0.7
2-H1’	3-H8	5.0	2.0	1.0
2-H2’	3-H8	2.5	0.7	0.5
2-H3’	3-H8	3.8	2.0	0.7
2-H8	3-H8	6.0	2.0	1.0
A3				
3-H2	3-H1’	5.0	2.0	1.0
3-H8	3-H1’	5.0	2.0	1.0
3-H8	3-H2’	5.0	2.0	1.0
3-H8	3-H3’	3.8	2.0	0.7
3-H8	3-H4’	5.0	2.0	1.0
3-H8	3-H5”	3.8	2.0	0.7
3-H8	3-H5’	3.8	2.0	0.7
3-H1’	4-H8	5.0	2.0	1.0
3-H2’	4-H1’	5.0	2.0	1.0
3-H2’	4-H8	2.5	0.7	0.5
3-H2	4-H1’	3.8	2.0	0.7
3-H2	4-H2’	6.0	2.0	1.0
3-H2	4-H2	5.0	2.0	1.0
3-H3’	4-H8	3.8	2.0	0.7
3-H8	4-H8	5.0	2.0	1.0
3-H2	25-H1’	6.0	2.0	1.0
3-H2	25-H2’	6.0	2.0	1.0
3-H2	26-H1’	3.8	2.0	0.7
3-H2	26-H2’	6.0	2.0	1.0
A4				
4-H2	4-H1’	5.0	2.0	1.0
4-H2	4-H2’	6.0	2.0	1.0
4-H8	4-H1’	5.0	2.0	1.0
4-H8	4-H2’	5.0	2.0	1.0
4-H8	4-H3’	3.8	2.0	0.7

4-H8	4-H4’	5.0	2.0	1.0
4-H8	4-H5”	3.8	2.0	0.7
4-H8	4-H5’	3.8	2.0	0.7
4-H1’	5-H1’	6.0	2.0	1.0
4-H1’	5-H5	6.0	2.0	1.0
4-H1’	5-H6	5.0	2.0	1.0
4-H2’	5-H1’	5.0	2.0	1.0
4-H2’	5-H5	3.8	2.0	0.7
4-H2’	5-H6	2.5	0.7	0.5
4-H2	5-H1’	3.8	2.0	0.7
4-H3’	5-H5	3.8	2.0	0.7
4-H3’	5-H6	3.8	2.0	0.7
4-H5”	5-H6	6.0	2.0	1.0
4-H8	5-H5	5.0	2.0	1.0
4-H8	5-H6	5.0	2.0	1.0
4-H2	24-H1’	6.0	2.0	1.0
4-H2	25-H1’	3.8	2.0	0.7
U5				
5-H2’	5-H5	5.0	2.0	1.0
5-H5	5-H3’	5.0	2.0	1.0
5-H6	5-H1’	5.0	2.0	1.0
5-H6	5-H2’	3.8	2.0	0.7
5-H6	5-H3’	3.8	2.0	0.7
5-H6	5-H4’	5.0	2.0	1.0
5-H6	5-H5”	3.8	2.0	0.7
5-H6	5-H5’	3.8	2.0	0.7
5-H1’	6-H8	5.0	2.0	1.0
5-H2’	6-H8	2.5	0.7	0.5
5-H3’	6-H8	3.8	2.0	0.7
5-H6	6-H2	6.0	2.0	1.0
5-H6	6-H8	6.0	2.0	1.0
5-H2’	21-H2	6.0	2.0	1.0
A6				
6-H1’	5-H2’	5.0	2.0	1.0
6-H2	6-H1’	5.0	2.0	1.0
6-H2	6-H2’	6.0	2.0	1.0
6-H8	6-H1’	3.8	2.0	0.7
6-H8	6-H2’	2.5	0.7	0.5
6-H8	6-H3’	3.8	2.0	0.7
6-H8	6-H4’	3.8	2.0	0.7
6-H8	6-H5”	3.8	2.0	0.7
6-H8	6-H5’	3.8	2.0	0.7
6-H1’	7-H2’	6.0	2.0	1.0
6-H1’	7-H3’	6.0	2.0	1.0
6-H1’	7-H1’	5.0	2.0	1.0
6-H1’	7-H5	5.0	2.0	1.0
6-H1’	7-H6	5.0	2.0	1.0
6-H2’	7-H1’	5.0	2.0	1.0
6-H2’	7-H5	3.8	2.0	0.7
6-H2’	7-H6	3.8	2.0	0.7
6-H2	7-H1’	3.8	2.0	0.7
6-H3’	7-H5	3.8	2.0	0.7
6-H3’	7-H6	3.8	2.0	0.7
6-H8	7-H5	6.0	2.0	1.0
6-H8	7-H6	5.0	2.0	1.0
6-H1’	8-H8	6.0	2.0	1.0
6-H1’	21-H2	5.0	2.0	1.0
6-H2’	21-H2	3.8	2.0	0.7
6-H2	21-H1’	5.0	2.0	1.0
6-H2	21-H2	5.0	2.0	1.0
6-H2	21-H2’	5.0	2.0	1.0
6-H2	21-H8	6.0	2.0	1.0
6-H3’	21-H2	6.0	2.0	1.0
6-H8	21-H2	5.0	2.0	1.0
6-H1’	22-H2	5.0	2.0	1.0
6-H2	22-H1’	6.0	2.0	1.0
6-H2	22-H2	3.8	2.0	0.7
6-H2	22-H8	5.0	2.0	1.0
6-H8	22-H2	5.0	2.0	1.0
6-H2	23-H1’	6.0	2.0	1.0
6-H2	23-H8	6.0	2.0	1.0
U7				
7-H5	7-H2’	5.0	2.0	1.0
7-H5	7-H3’	5.0	2.0	1.0
7-H5	7-H5”	6.0	2.0	1.0
7-H6	7-H1’	3.8	2.0	0.7
7-H6	7-H2’	5.0	2.0	1.0
7-H6	7-H3’	3.8	2.0	0.7
7-H6	7-H4’	5.0	2.0	1.0
7-H6	7-H5”	3.8	2.0	0.7
7-H6	7-H5’	3.8	2.0	0.7
7-H1’	8-H8	5.0	2.0	1.0
7-H2’	8-H1’	5.0	2.0	1.0
7-H2’	8-H8	2.5	0.7	0.5
7-H3’	8-H8	3.8	2.0	0.7
7-H6	8-H8	5.0	2.0	1.0
7-H1’	21-H2	5.0	2.0	1.0
7-H2’	21-H2	6.0	2.0	1.0
7-H5	21-H2	5.0	2.0	1.0
7-H6	21-H2	5.0	2.0	1.0
7-H1’	22-H2	5.0	2.0	1.0
G8				
8-H8	8-H1’	5.0	2.0	1.0
8-H8	8-H2’	5.0	2.0	1.0
8-H8	8-H3’	3.8	2.0	0.7
8-H8	8-H4’	5.0	2.0	1.0
8-H8	8-H5”	5.0	2.0	1.0
8-H1’	9-H5	6.0	2.0	1.0
8-H1’	9-H6	5.0	2.0	1.0
8-H2’	9-H1’	5.0	2.0	1.0
8-H2’	9-H5	3.8	2.0	0.7
8-H2’	9-H6	2.5	0.7	0.5
8-H3’	9-H5	3.8	2.0	0.7
8-H3’	9-H6	3.8	2.0	0.7
C9				
9-H5	9-H2’	5.0	2.0	1.0
9-H5	9-H3’	5.0	2.0	1.0
9-H5	9-H5”	6.0	2.0	1.0
9-H6	9-H1’	5.0	2.0	1.0
9-H6	9-H2’	5.0	2.0	1.0
9-H6	9-H3’	3.8	2.0	0.7
9-H6	9-H4’	5.0	2.0	1.0
9-H6	9-H5”	3.8	2.0	0.7
9-H1’	10-H6	5.0	2.0	1.0
9-H2’	10-H1’	5.0	2.0	1.0
9-H2’	10-H5	3.8	2.0	0.7
9-H2’	10-H6	2.5	0.7	0.5
9-H3’	10-H5	3.8	2.0	0.7
9-H3’	10-H6	3.8	2.0	0.7
9-H5	10-H5	5.0	2.0	1.0
9-H6	10-H5	5.0	2.0	1.0
9-H6	10-H6	5.0	2.0	1.0
U10				
10-H5	10-H1’	6.0	2.0	1.0
10-H5	10-H2’	5.0	2.0	1.0
10-H5	10-H3’	5.0	2.0	1.0
10-H5	10-H5”	6.0	2.0	1.0
10-H6	10-H1’	5.0	2.0	1.0
10-H6	10-H2’	5.0	2.0	1.0
10-H6	10-H3’	3.8	2.0	0.7
10-H6	10-H4’	5.0	2.0	1.0
10-H6	10-H5”	3.8	2.0	0.7
10-H1’	11-H6	3.8	2.0	0.7
10-H2’	11-H1’	5.0	2.0	1.0
10-H2’	11-H5	3.8	2.0	0.7
10-H2’	11-H6	2.5	0.7	0.5
10-H3’	11-H5	3.8	2.0	0.7
10-H3’	11-H6	3.8	2.0	0.7
10-H5	11-H5	5.0	2.0	1.0
10-H5	11-H6	6.0	2.0	1.0
10-H6	11-H6	5.0	2.0	1.0
C11				
11-H5	11-H2’	5.0	2.0	1.0
11-H5	11-H3’	5.0	2.0	1.0
11-H6	11-H1’	3.8	2.0	0.7
11-H6	11-H2’	5.0	2.0	1.0
11-H6	11-H3’	3.8	2.0	0.7
11-H6	11-H4’	5.0	2.0	1.0
11-H6	11-H5”	5.0	2.0	1.0
11-H6	11-H5’	5.0	2.0	1.0
11-H1’	12-H6	5.0	2.0	1.0
11-H2’	12-H1’	5.0	2.0	1.0
11-H2’	12-H5	3.8	2.0	0.7
11-H2’	12-H6	2.5	0.7	0.5
11-H3’	12-H5	3.8	2.0	0.7
11-H5	12-H5	3.8	2.0	0.7
11-H1’	17-H2	3.8	2.0	0.7
U12				
12-H6	11-H3’	3.8	2.0	0.7
12-H6	11-H5	6.0	2.0	1.0
12-H5	12-H2’	6.0	2.0	1.0
12-H5	12-H3’	5.0	2.0	1.0
12-H5	12-H5”	6.0	2.0	1.0
12-H6	12-H1’	5.0	2.0	1.0
12-H6	12-H2’	5.0	2.0	1.0
12-H6	12-H3’	3.8	2.0	0.7
12-H6	12-H4’	5.0	2.0	1.0
12-H6	12-H5”	3.8	2.0	0.7
12-H1’	13-H6	6.0	2.0	1.0
12-H2’	13-H2’	5.0	2.0	1.0
12-H2’	13-H3’	5.0	2.0	1.0
12-H2’	13-H5	6.0	2.0	1.0
12-H2’	13-H5”	6.0	2.0	1.0
12-H2’	13-H5’	5.0	2.0	1.0

12-H2’	13-H6	5.0	2.0	1.0
12-H4’	13-H5	2.5	0.7	0.5
12-H4’	13-H6	5.0	2.0	1.0
12-H5”	13-H5	5.0	2.0	1.0
12-H1’	14-H5	5.0	2.0	1.0
12-H2’	14-H1’	6.0	2.0	1.0
12-H2’	14-H2’	6.0	2.0	1.0
12-H2’	14-H5	2.5	0.7	0.5
12-H2’	14-H6	3.8	2.0	0.7
12-H3’	14-H5	3.8	2.0	0.7
12-H6	14-H5	6.0	2.0	1.0
U13				
13-H5	13-H1’	6.0	2.0	1.0
13-H5	13-H2’	5.0	2.0	1.0
13-H5	13-H3’	6.0	2.0	1.0
13-H6	13-H1’	3.8	2.0	0.7
13-H6	13-H2’	2.5	0.7	0.5
13-H6	13-H3’	3.8	2.0	0.7
13-H6	13-H4’	5.0	2.0	1.0
13-H6	13-H5”	5.0	2.0	1.0
13-H6	13-H5’	5.0	2.0	1.0
13-H1’	14-H5’	6.0	2.0	1.0
13-H1’	14-H5”	6.0	2.0	1.0
13-H1’	14-H6	5.0	2.0	1.0
13-H2’	14-H2’	5.0	2.0	1.0
13-H2’	14-H5	5.0	2.0	1.0
13-H2’	14-H6	3.8	2.0	0.7
13-H3’	14-H4’	5.0	2.0	1.0
13-H3’	14-H5”	3.8	2.0	0.7
13-H3’	14-H5’	5.0	2.0	1.0
13-H3’	14-H1’	6.0	2.0	1.0
13-H3’	14-H5	2.5	0.7	0.5
13-H3’	14-H6	2.5	0.7	0.5
13-H4’	14-H5	3.8	2.0	0.7
13-H4’	14-H6	5.0	2.0	1.0
13-H5”	14-H5	3.8	2.0	0.7
13-H5”	14-H6	5.0	2.0	1.0
13-H5’	14-H5	3.8	2.0	0.7
13-H5’	14-H6	5.0	2.0	1.0
13-H6	14-H5	6.0	2.0	1.0
13-H6	14-H6	6.0	2.0	1.0
C14				
14-H1’	14-H5	6.0	2.0	1.0
14-H5	14-H2’	3.8	2.0	0.7
14-H5	14-H3’	6.0	2.0	1.0
14-H5	14-H5”	6.0	2.0	1.0
14-H6	14-H1’	3.8	2.0	0.7
14-H6	14-H2’	2.5	0.7	0.5
14-H6	14-H3’	3.8	2.0	0.7
14-H6	14-H4’	5.0	2.0	1.0
14-H6	14-H5”	3.8	2.0	0.7
14-H6	14-H5’	3.8	2.0	0.7
14-H1’	15-H5’	5.0	2.0	1.0
14-H1’	15-H5”	5.0	2.0	1.0
14-H4’	15-H1’	3.8	2.0	0.7
14-H4’	15-H2’	5.0	2.0	1.0
14-H4’	15-H3’	3.8	2.0	0.7
14-H4’	15-H4’	3.8	2.0	0.7
14-H4’	15-H5”	3.8	2.0	0.7
14-H4’	15-H5’	3.8	2.0	0.7
14-H4’	15-H8	5.0	2.0	1.0
14-H5”	15-H1’	5.0	2.0	1.0
14-H5”	15-H4’	6.0	2.0	1.0
14-H5”	15-H5”	6.0	2.0	1.0
14-H5”	15-H5’	5.0	2.0	1.0
14-H5”	15-H8	5.0	2.0	1.0
14-H5’	15-H1’	5.0	2.0	1.0
14-H5’	15-H4’	6.0	2.0	1.0
14-H5’	15-H5”	6.0	2.0	1.0
14-H5’	15-H5’	5.0	2.0	1.0
14-H5’	15-H8	6.0	2.0	1.0
G15				
15-H8	15-H1’	2.5	0.7	0.5
15-H8	15-H2’	3.8	2.0	0.7
15-H8	15-H3’	6.0	2.0	1.0
15-H8	15-H4’	6.0	2.0	1.0
15-H8	15-H5”	6.0	2.0	1.0
15-H8	15-H5’	6.0	2.0	1.0
15-H1’	16-H1’	6.0	2.0	1.0
15-H1’	16-H4’	6.0	2.0	1.0
15-H2’	16-H1’	3.8	2.0	0.7
15-H2’	16-H2’	6.0	2.0	1.0
15-H2’	16-H3’	5.0	2.0	1.0
15-H2’	16-H4’	3.8	2.0	0.7
15-H2’	16-H5’	3.8	2.0	0.7
15-H2’	16-H5”	3.8	2.0	0.7
15-H2’	16-H8	3.8	2.0	0.7
15-H3’	16-H1’	5.0	2.0	1.0
15-H3’	16-H3’	6.0	2.0	1.0
15-H3’	16-H5’	5.0	2.0	1.0
15-H3’	16-H5”	5.0	2.0	1.0
15-H3’	16-H8	3.8	2.0	0.7
15-H8	16-H1’	6.0	2.0	1.0
G16				
16-H8	16-H1’	3.8	2.0	0.7
16-H8	16-H2’	3.8	2.0	0.7
16-H8	16-H3’	2.5	0.7	0.5
16-H8	16-H4’	3.8	2.0	0.7
16-H8	16-H5”	3.8	2.0	0.7
16-H8	16-H5’	3.8	2.0	0.7
16-H1’	17-H1’	6.0	2.0	1.0
16-H1’	17-H8	5.0	2.0	1.0
16-H2’	17-H1’	5.0	2.0	1.0
16-H2’	17-H8	2.5	0.7	0.5
16-H3’	17-H8	3.8	2.0	0.7
16-H8	17-H8	5.0	2.0	1.0
A17				
17-H2	17-H1’	5.0	2.0	1.0
17-H2	17-H2’	6.0	2.0	1.0
17-H8	17-H1’	5.0	2.0	1.0
17-H8	17-H2’	5.0	2.0	1.0
17-H8	17-H3’	3.8	2.0	0.7
17-H8	17-H4’	5.0	2.0	1.0
17-H8	17-H5”	3.8	2.0	0.7
17-H8	17-H5’	3.8	2.0	0.7
17-H1’	18-H8	5.0	2.0	1.0
17-H2’	18-H1’	5.0	2.0	1.0
17-H2’	18-H8	2.5	0.7	0.5
17-H2	18-H1’	3.8	2.0	0.7
17-H2	18-H2’	5.0	2.0	1.0
17-H2	18-H8	6.0	2.0	1.0
17-H3’	18-H8	3.8	2.0	0.7
17-H8	18-H8	5.0	2.0	1.0
G18				
18-H8	18-H1’	5.0	2.0	1.0
18-H8	18-H2’	3.8	2.0	0.7
18-H8	18-H3’	3.8	2.0	0.7
18-H8	18-H4’	5.0	2.0	1.0
18-H8	18-H5”	3.8	2.0	0.7
18-H8	18-H5’	5.0	2.0	1.0
18-H1’	19-H5	6.0	2.0	1.0
18-H1’	19-H6	5.0	2.0	1.0
18-H2’	19-H1’	5.0	2.0	1.0
18-H2’	19-H5	3.8	2.0	0.7
18-H2’	19-H6	2.5	0.7	0.5
18-H3’	19-H5	3.8	2.0	0.7
18-H3’	19-H6	3.8	2.0	0.7
18-H8	19-H5	5.0	2.0	1.0
18-H8	19-H6	5.0	2.0	1.0
C19				
19-H5	19-H2’	5.0	2.0	1.0
19-H5	19-H3’	5.0	2.0	1.0
19-H5	19-H5”	6.0	2.0	1.0
19-H6	19-H1’	3.8	2.0	0.7
19-H6	19-H2’	5.0	2.0	1.0
19-H6	19-H3’	2.5	0.7	0.5
19-H6	19-H4’	5.0	2.0	1.0
19-H6	19-H5”	3.8	2.0	0.7
19-H6	19-H5’	3.8	2.0	0.7
19-H1’	20-H6	5.0	2.0	1.0
19-H2’	20-H1’	5.0	2.0	1.0
19-H2’	20-H5	3.8	2.0	0.7
19-H2’	20-H6	2.5	0.7	0.5
19-H3’	20-H5	3.8	2.0	0.7
19-H3’	20-H6	3.8	2.0	0.7
19-H5	20-H5	5.0	2.0	1.0
19-H5	20-H6	5.0	2.0	1.0
19-H6	20-H6	6.0	2.0	1.0
U20				
20-H5	20-H2’	5.0	2.0	1.0
20-H5	20-H3’	5.0	2.0	1.0
20-H6	20-H1’	5.0	2.0	1.0
20-H6	20-H2’	3.8	2.0	0.7
20-H6	20-H3’	5.0	2.0	1.0
20-H6	20-H4’	3.8	2.0	0.7
20-H6	20-H5”	3.8	2.0	0.7
20-H1’	21-H8	5.0	2.0	1.0
20-H2’	21-H1’	6.0	2.0	1.0
20-H2’	21-H8	2.5	0.7	0.5
20-H3’	21-H1’	6.0	2.0	1.0
20-H3’	21-H8	3.8	2.0	0.7
20-H6	21-H8	5.0	2.0	1.0
A21				
21-H2	21-H1’	6.0	2.0	1.0

21-H2	21-H2’	6.0	2.0	1.0
21-H8	21-H1’	3.8	2.0	0.7
21-H8	21-H2’	2.5	0.7	0.5
21-H8	21-H3’	3.8	2.0	0.7
21-H8	21-H4’	5.0	2.0	1.0
21-H8	21-H5”	3.8	2.0	0.7
21-H8	21-H5’	3.8	2.0	0.7
A22				
21-H1’	22-H1’	6.0	2.0	1.0
21-H1’	22-H8	5.0	2.0	1.0
21-H2’	22-H1’	5.0	2.0	1.0
21-H2’	22-H8	3.8	2.0	0.7
21-H2	22-H1’	5.0	2.0	1.0
21-H2	22-H2	6.0	2.0	1.0
21-H3’	22-H8	3.8	2.0	0.7
21-H4’	22-H8	6.0	2.0	1.0
21-H5”	22-H8	6.0	2.0	1.0
21-H5’	22-H8	6.0	2.0	1.0
21-H8	22-H8	5.0	2.0	1.0
22-H2	22-H1’	5.0	2.0	1.0
22-H2	22-H2’	6.0	2.0	1.0
22-H8	22-H1’	3.8	2.0	0.7
22-H8	22-H2’	2.5	0.7	0.5
22-H8	22-H3’	3.8	2.0	0.7
22-H8	22-H4’	5.0	2.0	1.0
22-H8	22-H5”	5.0	2.0	1.0
22-H8	22-H5’	5.0	2.0	1.0
G23				
22-H1’	23-H1’	6.0	2.0	1.0
22-H1’	23-H8	5.0	2.0	1.0
22-H2’	23-H1’	6.0	2.0	1.0
22-H2’	23-H8	3.8	2.0	0.7
22-H2	23-H1’	3.8	2.0	0.7
22-H3’	23-H8	3.8	2.0	0.7
22-H4’	23-H8	5.0	2.0	1.0
22-H5”	23-H8	6.0	2.0	1.0
22-H8	23-H8	5.0	2.0	1.0
23-H8	23-H1’	3.8	2.0	0.7
23-H8	23-H2’	3.8	2.0	0.7
23-H8	23-H3’	3.8	2.0	0.7
23-H8	23-H4’	5.0	2.0	1.0
23-H8	23-H5”	3.8	2.0	0.7
23-H8	23-H5’	3.8	2.0	0.7
23-H1’	24-H5	6.0	2.0	1.0
23-H1’	24-H6	5.0	2.0	1.0
23-H2’	24-H1’	5.0	2.0	1.0
23-H2’	24-H5	3.8	2.0	0.7
23-H2’	24-H6	2.5	0.7	0.5
23-H3’	24-H5	3.8	2.0	0.7
23-H3’	24-H6	3.8	2.0	0.7
23-H4’	24-H5	6.0	2.0	1.0
23-H4’	24-H6	5.0	2.0	1.0
23-H8	24-H5	5.0	2.0	1.0
23-H8	24-H6	5.0	2.0	1.0
U24				
24-H5	24-H2’	5.0	2.0	1.0
24-H5	24-H3’	5.0	2.0	1.0
24-H5	24-H5”	6.0	2.0	1.0
24-H6	24-H1’	5.0	2.0	1.0
24-H6	24-H2’	5.0	2.0	1.0
24-H6	24-H3’	3.8	2.0	0.7
24-H6	24-H4’	5.0	2.0	1.0
24-H6	24-H5”	3.8	2.0	0.7
24-H1’	25-H1’	6.0	2.0	1.0
24-H2’	25-H1’	5.0	2.0	1.0
24-H2’	25-H5	3.8	2.0	0.7
24-H2’	25-H6	2.5	0.7	0.5
24-H3’	25-H5	3.8	2.0	0.7
24-H3’	25-H6	3.8	2.0	0.7
24-H5	25-H5	5.0	2.0	1.0
24-H6	25-H5	5.0	2.0	1.0
24-H6	25-H6	5.0	2.0	1.0
U25				
25-H5	25-H2’	5.0	2.0	1.0
25-H5	25-H3’	5.0	2.0	1.0
25-H6	25-H1’	3.8	2.0	0.7
25-H6	25-H2’	3.8	2.0	0.7
25-H6	25-H3’	2.5	0.7	0.5
25-H6	25-H4’	5.0	2.0	1.0
25-H6	25-H5”	3.8	2.0	0.7
25-H1’	26-H1’	6.0	2.0	1.0
25-H1’	26-H6	5.0	2.0	1.0
25-H2’	26-H1’	5.0	2.0	1.0
25-H2’	26-H5	3.8	2.0	0.7
25-H2’	26-H6	2.5	0.7	0.5
25-H3’	26-H5	3.8	2.0	0.7
25-H3’	26-H6	2.5	0.7	0.5
25-H5	26-H5	5.0	2.0	1.0
25-H6	26-H5	5.0	2.0	1.0
25-H6	26-H6	5.0	2.0	1.0
C26				
26-H5	26-H2’	5.0	2.0	1.0
26-H5	26-H3’	5.0	2.0	1.0
26-H5	26-H5”	6.0	2.0	1.0
26-H6	26-H1’	3.8	2.0	0.7
26-H6	26-H2’	5.0	2.0	1.0
26-H6	26-H3’	3.8	2.0	0.7
26-H6	26-H4’	5.0	2.0	1.0
26-H6	26-H5”	3.8	2.0	0.7
26-H6	26-H5’	3.8	2.0	0.7
C27				
26-H1’	27-H1’	6.0	2.0	1.0
26-H1’	27-H6	5.0	2.0	1.0
26-H2’	27-H1’	5.0	2.0	1.0
26-H2’	27-H5	3.8	2.0	0.7
26-H2’	27-H6	2.5	0.7	0.5
26-H3’	27-H5	3.8	2.0	0.7
26-H3’	27-H6	3.8	2.0	0.7
26-H5	27-H5	5.0	2.0	1.0
26-H6	27-H5	5.0	2.0	1.0
26-H6	27-H6	5.0	2.0	1.0
27-H5	27-H1’	6.0	2.0	1.0
27-H5	27-H2’	5.0	2.0	1.0
27-H5	27-H3’	5.0	2.0	1.0
27-H6	27-H1’	3.8	2.0	0.7
27-H6	27-H2’	3.8	2.0	0.7
27-H6	27-H3’	3.8	2.0	0.7
27-H6	27-H4’	5.0	2.0	1.0
27-H6	27-H5”	3.8	2.0	0.7
27-H6	27-H5’	3.8	2.0	0.7

Exchangeable protons:

G1				
1-H1’	1-H1	6.0	3.0	1.0
1-H1	2-H1	5.0	3.0	1.0
1-H1	2-H1’	5.0	3.0	1.0
1-H1	26-H41	5.0	2.0	1.0
1-H1	27-H1’	5.0	2.0	1.0
1-H1	27-H41	2.5	0.7	0.5
1-H1	27-H42	3.8	2.0	0.7
1-H1	27-H5	5.0	2.0	1.0
G2				
2-H2’	2-H1	6.0	3.0	1.0
2-H1	3-H2	3.8	2.0	0.7
2-H1	3-H62	6.0	3.0	1.0
2-H1	25-H3	3.8	2.0	0.7
2-H1	25-H5	6.0	2.0	1.0
2-H1	26-H1’	5.0	2.0	1.0
2-H1	26-H2’	6.0	2.0	1.0
2-H1	26-H41	2.5	0.7	0.5
2-H1	26-H42	3.8	2.0	0.7
2-H1	26-H5	5.0	2.0	1.0
2-H1	27-H5	6.0	2.0	1.0
2-H1	27-H6	6.0	2.0	1.0
A3				
3-H2	25-H3	2.5	0.7	0.5
3-H61	25-H3	2.5	0.7	0.5
3-H62	25-H3	3.8	2.0	0.7
A4				
4-H2	5-H3	6.0	3.0	1.0
4-H62	5-H3	5.0	2.0	1.0
4-H2	23-H1	5.0	2.0	1.0
4-H62	23-H1	5.0	2.0	1.0
4-H2	24-H3	2.5	0.7	0.5
4-H61	24-H3	2.5	0.7	0.5
4-H62	24-H3	3.8	2.0	0.7
4-H1’	25-H3	6.0	3.0	1.0
4-H2	25-H3	5.0	2.0	1.0
4-H61	25-H3	5.0	2.0	1.0
U5				
5-H3	6-H8	6.0	3.0	1.0
5-H3	22-H2	6.0	2.0	1.0
5-H3	23-H1	3.8	2.0	0.7
5-H3	23-H21	3.8	2.0	0.7
A6				
6-H1’	23-H1	6.0	3.0	1.0
G8				
8-H1	9-H1’	6.0	2.0	1.0
8-H1	9-H41	5.0	2.0	1.0
8-H1	9-H5	6.0	3.0	1.0
8-H1	18-H1	3.8	2.0	0.7
8-H1	18-H8	6.0	3.0	1.0
8-H1	19-H1’	5.0	2.0	1.0
8-H1	19-H41	2.5	0.7	0.5

8-H1	19-H42	3.8	2.0	0.7
8-H1	19-H5	5.0	2.0	1.0
8-H1	19-H6	6.0	3.0	1.0
8-H1	20-H5	6.0	2.0	1.0
8-H1	20-H6	6.0	3.0	1.0
C9				
9-H41	10-H3	5.0	2.0	1.0
9-H42	10-H3	5.0	2.0	1.0
9-H41	18-H1	2.5	0.7	0.5
9-H42	18-H1	3.8	2.0	0.7
9-H5	18-H1	5.0	2.0	1.0
9-H6	18-H1	6.0	3.0	1.0
9-H41	19-H41	5.0	2.0	1.0
U10				
10-H3	11-H41	5.0	3.0	1.0
10-H3	11-H42	5.0	2.0	1.0
10-H3	11-H5	5.0	2.0	1.0
10-H3	16-H1	5.0	2.0	1.0
10-H3	17-H2	2.5	0.7	0.5
10-H3	17-H61	2.5	0.7	0.5
10-H3	17-H62	3.8	2.0	0.7
10-H3	18-H1	3.8	2.0	0.7
10-H6	18-H1	6.0	3.0	1.0
C11				
11-H41	12-H3	5.0	3.0	1.0
11-H42	12-H3	5.0	2.0	1.0
11-H5	12-H3	5.0	2.0	1.0
11-H2’	16-H1	5.0	2.0	1.0
11-H41	16-H1	2.5	0.7	0.5
11-H42	16-H1	3.8	2.0	0.7
11-H5	16-H1	6.0	3.0	1.0
U12				
12-H1’	12-H3	5.0	2.0	1.0
12-H3	14-H1’	5.0	2.0	1.0
12-H3	14-H5	6.0	3.0	1.0
12-H1’	15-H1	3.8	2.0	0.7
12-H1’	15-H21	5.0	2.0	1.0
12-H2’	15-H1	3.8	2.0	0.7
12-H2’	15-H21	5.0	2.0	1.0
12-H3	15-H1	5.0	2.0	1.0
12-H3	15-H21	5.0	3.0	1.0
12-H1’	16-H1	3.8	2.0	0.7
12-H1’	16-H21	3.8	2.0	0.7
12-H1’	16-H22	3.8	2.0	0.7
12-H2’	16-H1	5.0	2.0	1.0
12-H3	16-H1	5.0	2.0	1.0
12-H5	16-H1	6.0	3.0	1.0
U13				
13-H2’	15-H1	5.0	2.0	1.0
13-H3’	15-H1	6.0	3.0	1.0
13-H6	15-H21	6.0	3.0	1.0
C14				
14-H1’	15-H1	3.8	2.0	0.7
14-H2’	15-H1	6.0	3.0	1.0
14-H5	15-H1	5.0	2.0	1.0
14-H5	15-H21	6.0	3.0	1.0
14-H6	15-H1	5.0	2.0	1.0
G15				
15-H1’	15-H1	6.0	3.0	1.0
15-H2’	15-H1	6.0	3.0	1.0
15-H1	16-H1	5.0	2.0	1.0
15-H1	16-H22	5.0	2.0	1.0
15-H1	16-H8	6.0	2.0	1.0
15-H21	16-H1	5.0	2.0	1.0
G16				
16-H2’	16-H1	6.0	2.0	1.0
16-H1	17-H1’	5.0	2.0	1.0
16-H1	17-H2	5.0	2.0	1.0
16-H1	17-H2’	6.0	2.0	1.0
16-H1	17-H8	5.0	2.0	1.0
16-H21	17-H1’	6.0	2.0	1.0
A17				
17-H2	18-H1	5.0	2.0	1.0
G18				
18-H1	19-H41	3.8	2.0	0.7
18-H1	19-H5	6.0	2.0	1.0
18-H1	19-H6	6.0	3.0	1.0
A22				
22-H2	23-H1	5.0	2.0	1.0
G23				
23-H1	24-H3	5.0	2.0	1.0
23-H1	24-H5	5.0	2.0	1.0
23-H21	24-H3	5.0	2.0	1.0
U24				
24-H3	25-H3	3.8	2.0	0.7
U25				
25-H1’	25-H3	5.0	2.0	1.0
25-H3	26-H41	5.0	2.0	1.0
25-H3	26-H42	5.0	2.0	1.0

Appendix 28. NOE-derived distance restraints used in D1kz structure calculations. Distance restraints derived from ^{1}H,^{1}H-NOESY spectra recorded in D_2O at 300 K (non-exchangeable protons) and in H_2O at 282 K (exchangeable protons), pD/pH 6.7, 60 mM KCl, 10 µM EDTA, 600 and 700 MHz:

Restraints are given in the following format (distances in Å):

<first proton> <second proton> <target distance> <lower limit> <upper limit>

NOE-derived distance restraints for non-exchangeable protons in the absence of $MgCl_2$ (excluding restraints in the κ region, which were only collected in the presence of Mg(II), see below):

G1				
1-H1’	1-H8	3.8	2.0	0.7
1-H2’	1-H8	5.0	2.0	1.0
1-H3’	1-H8	3.8	2.0	0.7
1-H4’	1-H8	5.0	2.0	1.0
1-H5’	1-H8	3.8	2.0	0.7
1-H5"	1-H8	3.8	2.0	0.7
1-H1’	2-H1’	6.0	2.0	1.0
1-H2’	2-H1’	5.0	2.0	1.0
1-H1’	2-H8	5.0	2.0	1.0
1-H2’	2-H8	2.5	0.7	0.5
1-H8	2-H8	5.0	2.0	1.0
G2				
2-H1’	2-H8	3.8	2.0	0.7
2-H2’	2-H8	5.0	2.0	1.0
2-H3’	2-H8	3.8	2.0	0.7
2-H4’	2-H8	5.0	2.0	1.0
2-H5’	2-H8	3.8	2.0	0.7
2-H5"	2-H8	3.8	2.0	0.7
2-H2’	3-H1’	5.0	2.0	1.0
2-H1’	3-H8	5.0	2.0	1.0
2-H2’	3-H8	2.5	0.7	0.5
2-H8	3-H8	5.0	2.0	1.0
A3				
3-H1’	3-H2	5.0	2.0	1.0
3-H2’	3-H2	5.0	2.0	1.0
3-H1’	3-H8	3.8	2.0	0.7
3-H2’	3-H8	5.0	2.0	1.0
3-H3’	3-H8	3.8	2.0	0.7
3-H4’	3-H8	5.0	2.0	1.0
3-H5’	3-H8	3.8	2.0	0.7
3-H5"	3-H8	3.8	2.0	0.7
3-H2’	4-H1’	5.0	2.0	1.0
3-H2	4-H1’	3.8	2.0	0.7
3-H2	4-H2’	6.0	2.0	1.0
3-H2	4-H2	5.0	2.0	1.0
3-H1’	4-H8	5.0	2.0	1.0
3-H2’	4-H8	2.5	0.7	0.5
3-H3’	4-H8	3.8	2.0	0.7

3-H2	4-H8	6.0	2.0	1.0
3-H8	4-H8	5.0	2.0	1.0
3-H2	47-H1’	5.0	2.0	1.0
3-H2	48-H1’	3.8	2.0	0.7
3-H2	48-H2’	6.0	2.0	1.0
3-H2	48-H6	6.0	2.0	1.0
A4				
4-H1’	4-H2	5.0	2.0	1.0
4-H1’	4-H8	3.8	2.0	0.7
4-H2’	4-H8	5.0	2.0	1.0
4-H3’	4-H8	3.8	2.0	0.7
4-H4’	4-H8	5.0	2.0	1.0
4-H5"	4-H8	3.8	2.0	0.7
4-H1’	5-H1’	6.0	2.0	1.0
4-H2’	5-H1’	5.0	2.0	1.0
4-H2	5-H1’	3.8	2.0	0.7
4-H2’	5-H2’	5.0	2.0	1.0
4-H2	5-H2’	6.0	2.0	2.0
4-H1’	5-H5	6.0	2.0	1.0
4-H2’	5-H5	3.8	2.0	1.0
4-H3’	5-H5	3.8	2.0	1.0
4-H8	5-H5	5.0	2.0	1.0
4-H1’	5-H6	5.0	2.0	1.0
4-H2’	5-H6	2.5	0.7	0.5
4-H3’	5-H6	3.8	2.0	1.0
4-H2	5-H6	5.0	2.0	1.0
4-H8	5-H6	5.0	2.0	1.0
4-H2	47-H1’	3.8	2.0	0.7
4-H2	47-H6	6.0	2.0	1.0
U5				
5-H1’	5-H5	6.0	2.0	1.0
5-H2’	5-H5	5.0	2.0	1.0
5-H3’	5-H5	5.0	2.0	1.0
5-H1’	5-H6	3.8	2.0	0.7
5-H2’	5-H6	5.0	2.0	1.0
5-H3’	5-H6	3.8	2.0	0.7
5-H4’	5-H6	5.0	2.0	1.0
5-H5’	5-H6	3.8	2.0	0.7
5-H5"	5-H6	3.8	2.0	0.7
5-H2’	6-H1’	5.0	2.0	1.0
5-H2’	6-H2’	5.0	2.0	1.0
5-H1’	6-H8	5.0	2.0	1.0
5-H2’	6-H8	2.5	0.7	0.5
5-H3’	6-H8	3.8	2.0	0.7
5-H6	6-H8	6.0	2.0	1.0
A6				
6-H1’	6-H2	5.0	2.0	1.0
6-H2’	6-H2	6.0	2.0	1.0
6-H1’	6-H8	3.8	2.0	0.7
6-H2’	6-H8	3.8	2.0	0.7
6-H3’	6-H8	3.8	2.0	0.7
6-H4’	6-H8	5.0	2.0	1.0
6-H5’	6-H8	3.8	2.0	1.0
6-H5"	6-H8	5.0	2.0	1.0
6-H1’	7-H1’	5.0	2.0	1.0
6-H2’	7-H1’	5.0	2.0	2.0
6-H2	7-H1’	3.8	2.0	1.0
6-H1’	7-H5	5.0	2.0	1.0
6-H2’	7-H5	3.8	2.0	0.7
6-H3’	7-H5	3.8	2.0	0.7
6-H1’	7-H6	3.8	2.0	1.0
6-H2’	7-H6	3.8	2.0	0.7
6-H3’	7-H6	3.8	2.0	0.7
6-H8	7-H5	6.0	2.0	1.0
6-H8	7-H6	5.0	2.0	1.0
6-H2	43-H1’	5.0	2.0	1.0
6-H2	43-H2’	5.0	2.0	1.0
6-H1’	43-H2	5.0	2.0	1.0
6-H2’	43-H2	3.8	2.0	0.7
6-H2	43-H2	3.8	2.0	0.7
6-H8	43-H2	6.0	2.0	1.0
6-H2	43-H8	6.0	2.0	1.0
6-H2	44-H1’	5.0	2.0	1.0
6-H1’	44-H2	5.0	2.0	1.0
6-H2	44-H2	3.8	2.0	0.7
6-H2	44-H8	6.0	2.0	1.0
6-H8	44-H2	6.0	2.0	1.0
6-H2	45-H1’	6.0	2.0	1.0
U7				
7-H2’	7-H5	5.0	2.0	1.0
7-H5"	7-H5	6.0	2.0	1.0
7-H1’	7-H6	3.8	2.0	0.7
7-H2’	7-H6	3.8	2.0	0.7
7-H5"	7-H6	3.8	2.0	0.7
7-H1’	8-H1’	6.0	2.0	1.0
7-H2’	8-H1’	5.0	2.0	2.0
7-H1’	8-H8	5.0	2.0	1.0
7-H2’	8-H8	2.5	0.7	0.5
7-H6	8-H8	5.0	2.0	1.0
7-H1’	43-H2	5.0	2.0	1.0
7-H2’	43-H2	6.0	2.0	1.0
7-H5	43-H2	5.0	2.0	1.0
7-H6	43-H2	5.0	2.0	1.0
G8				
8-H1’	8-H8	5.0	3.0	1.0
8-H2’	8-H8	3.8	2.0	0.7
8-H3’	8-H8	3.8	2.0	0.7
8-H8	8-H5"	3.8	2.0	0.7
8-H2’	9-H1’	5.0	2.0	1.0
8-H1’	9-H5	6.0	2.0	1.0
8-H2’	9-H5	3.8	2.0	0.7
8-H3’	9-H5	3.8	2.0	0.7
8-H1’	9-H6	5.0	2.0	1.0
8-H2’	9-H6	2.5	0.7	0.5
8-H3’	9-H6	3.8	2.0	0.7
8-H8	9-H5	6.0	2.0	1.0
C9				
9-H1’	9-H5	6.0	2.0	1.0
9-H2’	9-H5	5.0	2.0	1.0
9-H3’	9-H5	5.0	2.0	1.0
9-H1’	9-H6	3.8	2.0	0.7
9-H2’	9-H6	5.0	2.0	1.0
9-H3’	9-H6	3.8	2.0	0.7
9-H5"	9-H6	3.8	2.0	0.7
9-H2’	10-H1’	5.0	2.0	1.0
9-H1’	10-H5	5.0	2.0	1.0
9-H2’	10-H5	3.8	2.0	0.7
9-H3’	10-H5	5.0	2.0	1.0
9-H5	10-H5	5.0	2.0	1.0
9-H1’	10-H6	5.0	2.0	1.0
9-H2’	10-H6	2.5	0.7	0.5
9-H3’	10-H6	3.8	2.0	0.7
9-H6	10-H6	5.0	2.0	1.0
U10				
10-H2’	10-H5	5.0	2.0	1.0
10-H1’	10-H6	3.8	2.0	0.7
10-H2’	10-H6	5.0	2.0	1.0
…				
C14				
14-H2’	14-H5	5.0	2.0	1.0
14-H3’	14-H5	5.0	2.0	1.0
14-H5"	14-H5	6.0	2.0	1.0
14-H1’	14-H6	3.8	2.0	0.7
14-H2’	14-H6	3.8	2.0	0.7
14-H3’	14-H6	3.8	2.0	0.7
14-H4’	14-H6	5.0	2.0	1.0
14-H5’	14-H6	3.8	2.0	0.7
14-H5"	14-H6	3.8	2.0	0.7
14-H1’	15-H1’	5.0	2.0	1.0
14-H2’	15-H1’	5.0	2.0	1.0
14-H2’	15-H2’	5.0	2.0	1.0
14-H1’	15-H8	5.0	2.0	1.0
14-H2’	15-H8	2.5	0.7	0.5
14-H6	15-H8	5.0	2.0	1.0
G15				
15-H1’	15-H8	3.8	2.0	0.7
15-H2’	15-H8	3.8	2.0	0.7
15-H3’	15-H8	3.8	2.0	0.7
15-H4’	15-H8	5.0	2.0	1.0
15-H5"	15-H8	3.8	2.0	0.7
15-H2’	16-H1’	6.0	2.0	1.0
15-H1’	16-H8	6.0	2.0	1.0
15-H2’	16-H8	5.0	2.0	1.0
15-H3’	16-H8	5.0	2.0	1.0
15-H1’	17-H8	5.0	2.0	1.0
15-H2’	17-H8	5.0	2.0	1.0
15-H3’	17-H8	6.0	2.0	1.0
A16				
16-H1’	16-H2	5.0	2.0	1.0
16-H1’	16-H8	3.8	2.0	0.7
16-H2’	16-H8	3.8	2.0	0.7
16-H3’	16-H8	3.8	2.0	0.7
16-H4’	16-H8	5.0	2.0	1.0
16-H5’	16-H8	3.8	2.0	0.7
16-H5"	16-H8	5.0	3.0	1.0
16-H1’	17-H1’	5.0	2.0	1.0
16-H2’	17-H1’	3.8	2.0	0.7
16-H2	17-H1’	3.8	2.0	0.7
16-H1’	17-H8	5.0	2.0	1.0
16-H2’	17-H8	2.5	0.7	0.5
16-H3’	17-H8	2.5	0.7	0.5
16-H4’	17-H8	5.0	2.0	1.0
16-H8	17-H8	5.0	2.0	1.0
A17				

17-H1’	17-H8	3.8	2.0	0.7
17-H2’	17-H8	3.8	2.0	0.7
17-H3’	17-H8	3.8	2.0	0.7
17-H4’	17-H8	5.0	2.0	1.0
17-H5"	17-H8	3.8	2.0	0.7
17-H1’	18-H1’	6.0	2.0	1.0
17-H2’	18-H1’	5.0	2.0	1.0
17-H2	18-H1’	3.8	2.0	0.7
17-H2	18-H2	5.0	2.0	1.0
17-H1’	18-H8	5.0	2.0	1.0
17-H2’	18-H8	2.5	0.7	0.5
17-H5"	18-H8	5.0	2.0	1.0
17-H2	18-H8	5.0	2.0	1.0
17-H8	18-H8	5.0	2.0	1.0
A18				
18-H1’	18-H2	6.0	2.0	1.0
18-H1’	18-H8	3.8	2.0	0.7
18-H2’	18-H8	3.8	2.0	0.7
18-H3’	18-H8	3.8	2.0	0.7
18-H4’	18-H8	5.0	3.0	1.0
18-H5’	18-H8	3.8	2.0	0.7
18-H5"	18-H8	3.8	2.0	0.7
...				
C25				
25-H1’	25-H5	5.0	2.0	1.0
25-H2’	25-H5	5.0	2.0	1.0
25-H3’	25-H5	5.0	2.0	1.0
25-H1’	25-H6	3.8	2.0	0.7
25-H2’	25-H6	5.0	2.0	1.0
25-H3’	25-H6	3.8	2.0	0.7
25-H1’	26-H1’	6.0	2.0	1.0
25-H2’	26-H1’	5.0	2.0	1.0
25-H1’	26-H8	5.0	2.0	1.0
25-H2’	26-H8	2.5	0.7	0.5
25-H3’	26-H8	3.8	2.0	0.7
25-H6	26-H8	5.0	2.0	1.0
A26				
26-H1’	26-H2	6.0	2.0	1.0
26-H2’	26-H2	5.0	2.0	1.0
26-H1’	26-H8	5.0	2.0	1.0
26-H2’	26-H8	5.0	2.0	1.0
26-H3’	26-H8	3.8	2.0	0.7
26-H5"	26-H8	5.0	2.0	1.0
26-H1’	27-H1’	6.0	2.0	1.0
26-H2’	27-H1’	5.0	2.0	1.0
26-H2	27-H1’	3.8	2.0	0.7
26-H2	27-H2’	6.0	2.0	1.0
26-H1’	27-H8	5.0	2.0	1.0
26-H2’	27-H8	2.5	0.7	0.5
26-H2	27-H8	6.0	2.0	1.0
26-H3’	27-H8	3.8	2.0	0.7
26-H8	27-H8	5.0	2.0	1.0
26-H2	35-H1’	6.0	2.0	1.0
26-H2	36-H1’	3.8	2.0	0.7
26-H2	36-H2’	6.0	2.0	1.0
26-H2	36-H8	6.0	2.0	1.0
G27				
27-H1’	27-H8	5.0	2.0	1.0
27-H2’	27-H8	5.0	2.0	1.0
27-H3’	27-H8	3.8	2.0	0.7
27-H4’	27-H8	5.0	2.0	1.0
27-H5"	27-H8	5.0	2.0	1.0
27-H1’	28-H1’	6.0	2.0	1.0
27-H2’	28-H1’	5.0	2.0	1.0
27-H1’	28-H5	6.0	2.0	1.0
27-H2’	28-H5	3.8	2.0	0.7
27-H3’	28-H5	3.8	2.0	0.7
27-H8	28-H5	6.0	2.0	1.0
27-H1’	28-H6	5.0	2.0	1.0
27-H2’	28-H6	2.5	0.7	0.5
27-H3’	28-H6	3.8	2.0	0.7
27-H8	28-H6	5.0	2.0	1.0
C28				
28-H1’	28-H5	6.0	2.0	1.0
28-H2’	28-H5	5.0	2.0	1.0
28-H3’	28-H5	5.0	2.0	1.0
28-H1’	28-H6	3.8	2.0	0.7
28-H2’	28-H6	5.0	2.0	1.0
28-H3’	28-H6	3.8	2.0	0.7
28-H4’	28-H6	5.0	2.0	1.0
28-H5"	28-H6	3.8	2.0	0.7
28-H1’	29-H1’	6.0	2.0	1.0
28-H2’	29-H1’	5.0	2.0	1.0
28-H2’	29-H2’	5.0	2.0	1.0
28-H1’	29-H5	5.0	2.0	1.0
28-H2’	29-H5	3.8	2.0	0.7
28-H3’	29-H5	3.8	2.0	0.7
28-H5	29-H5	3.8	2.0	0.7
28-H6	29-H5	3.8	2.0	0.7
28-H1’	29-H6	5.0	2.0	1.0
28-H2’	29-H6	2.5	0.7	0.5
28-H3’	29-H6	3.8	2.0	0.7
28-H5	29-H6	5.0	2.0	1.0
28-H6	29-H6	5.0	2.0	1.0
U29				
29-H2’	29-H5	6.0	2.0	1.0
29-H3’	29-H5	5.0	2.0	1.0
29-H5"	29-H5	6.0	2.0	1.0
29-H1’	29-H6	3.8	2.0	0.7
29-H2’	29-H6	3.8	2.0	0.7
29-H3’	29-H6	3.8	2.0	0.7
29-H5"	29-H6	3.8	2.0	0.7
29-H2’	30-H1’	6.0	2.0	1.0
29-H1’	30-H2’	6.0	2.0	1.0
29-H2’	30-H2’	3.8	2.0	0.7
29-H1’	30-H3’	6.0	2.0	1.0
29-H2’	30-H3’	5.0	2.0	1.0
29-H2’	30-H5’	5.0	2.0	1.0
29-H1’	30-H5	6.0	2.0	1.0
29-H2’	30-H5	6.0	2.0	2.0
29-H4’	30-H5	3.8	2.0	0.7
29-H5"	30-H5	5.0	2.0	1.0
29-H1’	30-H6	5.0	2.0	1.0
29-H2’	30-H6	5.0	2.0	1.0
29-H4’	30-H6	5.0	2.0	1.0
29-H2’	31-H1’	6.0	2.0	1.0
29-H2’	31-H2’	6.0	2.0	1.0
29-H1’	31-H5	5.0	2.0	1.0
29-H2’	31-H5	2.5	0.7	0.5
29-H3’	31-H5	3.8	2.0	0.7
29-H4’	31-H5	6.0	2.0	1.0
29-H6	31-H5	6.0	2.0	1.0
29-H1’	31-H6	6.0	2.0	1.0
29-H2’	31-H6	3.8	2.0	0.7
U30				
30-H1’	30-H5	5.0	2.0	1.0
30-H2’	30-H5	5.0	2.0	1.0
30-H3’	30-H5	6.0	2.0	1.0
30-H5"	30-H5	6.0	2.0	1.0
30-H1’	30-H6	3.8	2.0	0.7
30-H2’	30-H6	2.5	0.7	0.5
30-H3’	30-H6	3.8	2.0	0.7
30-H5’	30-H6	3.8	2.0	0.7
30-H1’	31-H5’	6.0	2.0	1.0
30-H2’	31-H2’	6.0	2.0	1.0
30-H2’	31-H5	5.0	2.0	1.0
30-H3’	31-H2’	5.0	2.0	1.0
30-H3’	31-H4’	5.0	2.0	1.0
30-H3’	31-H5"	5.0	2.0	1.0
30-H3’	31-H5’	5.0	2.0	1.0
30-H3’	31-H5	3.8	2.0	0.7
30-H4’	31-H5	5.0	2.0	1.0
30-H5’	31-H5	3.8	2.0	0.7
30-H5"	31-H5	2.5	0.7	0.5
30-H6	31-H5	5.0	2.0	1.0
30-H1’	31-H6	6.0	2.0	1.0
30-H2’	31-H6	5.0	2.0	1.0
30-H3’	31-H6	2.5	0.7	0.5
30-H4’	31-H6	5.0	2.0	1.0
30-H5’	31-H6	5.0	2.0	1.0
30-H6	31-H6	5.0	2.0	1.0
C31				
31-H1’	31-H5	6.0	2.0	1.0
31-H2’	31-H5	5.0	2.0	1.0
31-H3’	31-H5	6.0	2.0	1.0
31-H5"	31-H5	6.0	2.0	1.0
31-H1’	31-H6	3.8	2.0	0.7
31-H2’	31-H6	2.5	0.7	0.5
31-H3’	31-H6	3.8	2.0	0.7
31-H4’	31-H6	5.0	2.0	1.0
31-H5"	31-H6	5.0	2.0	1.0
31-H5’	31-H6	5.0	2.0	1.0
31-H1’	32-H5’	5.0	2.0	1.0
31-H4’	32-H1’	3.8	2.0	0.7
31-H5’	32-H1’	5.0	2.0	1.0
31-H5"	32-H1’	6.0	2.0	1.0
31-H4’	32-H2’	5.0	2.0	1.0
31-H4’	32-H3’	5.0	2.0	1.0
31-H4’	32-H4’	5.0	2.0	1.0
31-H4’	32-H5’	3.8	2.0	0.7
31-H4’	32-H5"	3.8	2.0	0.7
31-H5’	32-H5’	5.0	2.0	2.0
31-H5’	32-H5"	6.0	2.0	2.0
31-H5"	32-H5’	5.0	2.0	2.0

31-H5"	32-H5"	6.0	2.0	2.0
31-H4’	32-H8	5.0	2.0	1.0
31-H5"	32-H8	5.0	2.0	2.0
G32				
32-H1’	32-H8	2.5	0.7	0.5
32-H2’	32-H8	3.8	2.0	0.7
32-H4’	32-H8	5.0	2.0	1.0
32-H5"	32-H8	6.0	2.0	2.0
32-H5’	32-H8	6.0	2.0	2.0
32-H1’	33-H1’	6.0	2.0	1.0
32-H2’	33-H1’	3.8	2.0	0.7
32-H3’	33-H1’	5.0	2.0	1.0
32-H8	33-H1’	6.0	2.0	1.0
32-H2’	33-H2’	6.0	2.0	1.0
32-H2’	33-H4’	3.8	2.0	0.7
32-H3’	33-H4’	5.0	2.0	1.0
32-H2’	33-H5’	3.8	2.0	0.7
32-H3’	33-H5’	5.0	2.0	1.0
32-H2’	33-H5"	3.8	2.0	0.7
32-H3’	33-H5"	5.0	2.0	1.0
32-H1’	33-H8	6.0	2.0	1.0
32-H2’	33-H8	3.8	2.0	0.7
32-H3’	33-H8	3.8	2.0	0.7
G33				
33-H1’	33-H8	3.8	2.0	0.7
33-H2’	33-H8	5.0	2.0	1.0
33-H3’	33-H8	3.8	2.0	0.7
33-H4’	33-H8	3.8	2.0	0.7
33-H5’	33-H8	3.8	2.0	0.7
33-H5"	33-H8	3.8	2.0	0.7
33-H1’	34-H1’	5.0	2.0	1.0
33-H2’	34-H1’	5.0	2.0	1.0
33-H1’	34-H5	5.0	2.0	1.0
33-H2’	34-H5	3.8	2.0	0.7
33-H3’	34-H5	3.8	2.0	0.7
33-H8	34-H5	5.0	2.0	1.0
33-H1’	34-H6	3.8	2.0	0.7
33-H2’	34-H6	2.5	0.7	0.5
33-H3’	34-H6	3.8	2.0	0.7
33-H8	34-H6	6.0	2.0	1.0
C34				
34-H1’	34-H5	5.0	2.0	1.0
34-H2’	34-H5	5.0	2.0	1.0
34-H3’	34-H5	5.0	2.0	1.0
34-H1’	34-H6	3.8	2.0	1.0
34-H2’	34-H6	5.0	2.0	1.0
34-H3’	34-H6	3.8	2.0	1.0
34-H5"	34-H6	3.8	2.0	1.0
34-H2’	35-H1’	5.0	2.0	1.0
34-H2’	35-H2’	5.0	2.0	1.0
34-H2’	35-H5	3.8	2.0	1.0
34-H3’	35-H5	3.8	2.0	1.0
34-H5	35-H5	5.0	2.0	1.0
34-H1’	35-H6	5.0	2.0	1.0
34-H2’	35-H6	2.5	0.7	0.5
34-H3’	35-H6	3.8	2.0	1.0
34-H6	35-H6	5.0	2.0	1.0
U35				
35-H1’	35-H6	3.8	2.0	0.7
35-H2’	35-H6	5.0	2.0	1.0
35-H3’	35-H6	3.8	2.0	0.7
35-H4’	35-H6	5.0	2.0	1.0
35-H5"	35-H6	3.8	2.0	0.7
…				
G40				
40-H2’	40-H8	5.0	2.0	1.0
40-H3’	40-H8	3.8	2.0	0.7
40-H2’	41-H1’	5.0	2.0	1.0
40-H2’	41-H5	3.8	2.0	0.7
40-H3’	41-H5	5.0	3.0	1.0
40-H2’	41-H6	2.5	0.7	0.5
40-H3’	41-H6	3.8	2.0	0.7
40-H8	41-H5	6.0	2.0	1.0
40-H8	41-H6	5.0	2.0	1.0
C41				
41-H1’	41-H5	5.0	2.0	1.0
41-H3’	41-H5	5.0	2.0	1.0
41-H1’	41-H6	5.0	3.0	1.0
41-H2’	41-H6	5.0	2.0	1.0
41-H3’	41-H6	3.8	2.0	0.7
41-H1’	42-H1’	6.0	2.0	1.0
41-H2’	42-H1’	5.0	2.0	1.0
41-H2’	42-H2’	5.0	2.0	1.0
41-H2’	42-H5	3.8	2.0	0.7
41-H3’	42-H5	5.0	2.0	1.0
41-H5	42-H5	5.0	2.0	1.0
41-H6	42-H5	6.0	2.0	1.0
41-H1’	42-H6	5.0	2.0	1.0
41-H2’	42-H6	2.5	0.7	0.5
41-H3’	42-H6	3.8	2.0	0.7
41-H5	42-H6	6.0	2.0	1.0
41-H6	42-H6	6.0	2.0	1.0
U42				
42-H2’	42-H5	5.0	2.0	1.0
42-H3’	42-H5	5.0	2.0	1.0
42-H1’	42-H6	3.8	2.0	0.7
42-H2’	42-H6	3.8	2.0	0.7
42-H3’	42-H6	5.0	3.0	1.0
42-H4’	42-H6	5.0	2.0	1.0
42-H2’	43-H2’	5.0	2.0	1.0
42-H2’	43-H3’	6.0	2.0	1.0
42-H1’	43-H8	5.0	2.0	1.0
42-H2’	43-H8	2.5	0.7	0.5
42-H3’	43-H8	3.8	2.0	0.7
42-H6	43-H8	5.0	2.0	1.0
A43				
43-H1’	43-H8	3.8	2.0	0.7
43-H2’	43-H8	2.5	0.7	0.5
43-H3’	43-H8	3.8	2.0	0.7
43-H4’	43-H8	5.0	2.0	1.0
43-H5’	43-H8	3.8	2.0	0.7
43-H5"	43-H8	5.0	2.0	1.0
43-H1’	44-H1’	6.0	2.0	1.0
43-H2’	44-H1’	6.0	2.0	3.0
43-H2	44-H1’	6.0	2.0	1.0
43-H2	44-H2	6.0	2.0	1.0
43-H1’	44-H8	5.0	2.0	1.0
43-H2’	44-H8	5.0	2.0	1.0
43-H3’	44-H8	5.0	3.0	1.0
43-H4’	44-H8	6.0	2.0	1.0
43-H5"	44-H8	6.0	2.0	1.0
43-H8	44-H8	6.0	2.0	1.0
A44				
44-H1’	44-H2	6.0	2.0	1.0
44-H2’	44-H2	6.0	3.0	1.0
44-H1’	44-H8	3.8	2.0	0.7
44-H2’	44-H8	3.8	2.0	0.7
44-H3’	44-H8	3.8	2.0	0.7
44-H4’	44-H8	5.0	2.0	1.0
44-H5’	44-H8	5.0	2.0	1.0
44-H5"	44-H8	5.0	2.0	1.0
44-H1’	45-H1’	6.0	2.0	1.0
44-H2’	45-H1’	6.0	2.0	3.0
44-H2	45-H1’	3.8	2.0	0.7
44-H1’	45-H8	5.0	2.0	1.0
44-H2’	45-H8	3.8	2.0	0.7
44-H3’	45-H8	5.0	2.0	1.0
44-H8	45-H8	5.0	2.0	1.0
G45				
45-H1’	45-H8	5.0	3.0	1.0
45-H2’	45-H8	5.0	2.0	1.0
45-H3’	45-H8	3.8	2.0	0.7
45-H2’	46-H1’	5.0	2.0	1.0
45-H2’	46-H5	3.8	2.0	0.7
45-H8	46-H5	5.0	2.0	1.0
45-H3’	46-H5	3.8	2.0	0.7
45-H1’	46-H6	5.0	2.0	1.0
45-H2’	46-H6	2.5	0.7	0.5
45-H3’	46-H6	5.0	3.0	1.0
U46				
46-H1’	46-H5	6.0	2.0	1.0
46-H2’	46-H5	5.0	2.0	1.0
46-H3’	46-H5	5.0	2.0	1.0
46-H1’	46-H6	3.8	2.0	0.7
46-H2’	46-H6	5.0	2.0	1.0
46-H3’	46-H6	3.8	2.0	0.7
46-H1’	47-H1’	6.0	2.0	1.0
46-H2’	47-H1’	5.0	2.0	1.0
46-H5	47-H1’	6.0	2.0	2.0
46-H2’	47-H5	3.8	2.0	0.7
46-H3’	47-H5	3.8	2.0	0.7
46-H5	47-H5	5.0	2.0	1.0
46-H1’	47-H6	5.0	2.0	1.0
46-H2’	47-H6	2.5	0.7	0.5
46-H3’	47-H6	3.8	2.0	0.7
46-H6	47-H6	5.0	2.0	1.0
U47				
47-H1’	47-H6	3.8	2.0	0.7
47-H2’	47-H6	5.0	2.0	1.0
47-H3’	47-H6	3.8	2.0	0.7
47-H5"	47-H6	3.8	2.0	0.7
47-H1’	48-H1’	6.0	2.0	1.0
47-H2’	48-H1’	5.0	2.0	1.0
47-H2’	48-H2’	6.0	2.0	1.0

47-H2’	48-H5	3.8	2.0	0.7
47-H3’	48-H5	3.8	2.0	0.7
47-H1’	48-H6	5.0	2.0	1.0
47-H2’	48-H6	2.5	0.7	0.5
47-H3’	48-H6	3.8	2.0	0.7
47-H6	48-H6	5.0	2.0	1.0
C48				
48-H1’	48-H5	5.0	2.0	1.0
48-H2’	48-H5	5.0	2.0	1.0
48-H3’	48-H5	5.0	2.0	1.0
48-H1’	48-H6	3.8	2.0	0.7
48-H2’	48-H6	5.0	2.0	1.0
48-H3’	48-H6	3.8	2.0	0.7
48-H5"	48-H6	5.0	2.0	1.0
48-H1’	49-H1’	5.0	2.0	1.0
48-H2’	49-H1’	5.0	2.0	1.0
48-H2’	49-H2’	5.0	2.0	1.0
48-H2’	49-H5	3.8	2.0	1.0
48-H3’	49-H5	3.8	2.0	1.0
48-H5	49-H5	5.0	2.0	1.0
48-H1’	49-H6	5.0	2.0	1.0
48-H2’	49-H6	2.5	0.7	0.5
48-H3’	49-H6	3.8	2.0	0.7
48-H6	49-H6	5.0	2.0	1.0
C49				
49-H1’	49-H5	6.0	2.0	1.0
49-H2’	49-H5	5.0	2.0	1.0
49-H3’	49-H5	5.0	2.0	1.0
49-H1’	49-H6	3.8	2.0	0.7
49-H2’	49-H6	3.8	2.0	0.7
49-H3’	49-H6	3.8	2.0	0.7
49-H4’	49-H6	5.0	2.0	1.0
49-H5"	49-H6	5.0	2.0	1.0

NOE-derived distance restraints for non-exchangeable protons in the κ region in the presence of 9 mM $MgCl_2$:

U10				
10-H1’	11-H1’	6.0	3.0	1.0
10-H2’	11-H1’	5.0	3.0	1.0
10-H5	11-H5	5.0	3.0	1.0
10-H1’	11-H6	5.0	2.0	1.0
10-H2’	11-H6	3.8	2.0	0.7
10-H6	11-H6	5.0	2.0	1.0
10-H1’	22-H2	5.0	2.0	1.0
10-H2’	22-H2	4.0	2.0	1.0
10-H1’	39-H2	6.0	3.0	1.0
C11				
11-H6	11-H1’	5.0	3.0	1.0
11-H6	11-H2’	3.8	2.0	0.7
11-H6	11-H4’	5.0	3.0	1.0
11-H4’	12-H1’	5.0	3.0	1.0
11-H1’	22-H2	5.0	3.0	1.0
11-H4’	22-H2	5.0	3.0	1.0
11-H1’	23-H1’	3.8	2.0	0.7
11-H1’	23-H2	6.0	2.0	1.0
11-H4’	23-H1’	5.0	3.0	1.0
11-H4’	23-H2	5.0	3.0	1.0
11-H1’	24-H6	6.0	3.0	2.0
11-H2’	24-H6	5.0	3.0	1.0
11-H1’	39-H2	6.0	3.0	1.0
A12				
12-H1’	12-H8	5.0	3.0	1.0
12-H1’	13-H8	5.0	3.0	1.0
12-H2’	13-H1’	5.0	2.0	1.0
12-H2’	13-H2’	6.0	3.0	1.0
12-H2’	13-H8	3.8	2.0	0.7
12-H2	13-H1’	3.8	2.0	0.7
12-H2	13-H2’	6.0	3.0	1.0
12-H1’	23-H2	3.8	2.0	0.7
12-H2’	23-H2	5.0	2.0	1.0
12-H8	23-H1’	5.0	3.0	1.0
12-H8	23-H2’	3.8	2.0	0.7
12-H8	23-H2	5.0	3.0	1.0
12-H8	23-H8	6.0	2.0	2.0
A13				
13-H8	13-H1’	5.0	2.0	1.0
13-H8	13-H2’	5.0	3.0	1.0
13-H1’	14-H1’	6.0	2.0	1.0
13-H2’	14-H1’	5.0	3.0	1.0
13-H1’	14-H5	5.0	3.0	1.0
13-H1’	14-H6	3.8	2.0	0.7
13-H2’	14-H6	2.5	0.7	0.5
13-H2	14-H1’	5.0	3.0	1.0
13-H2	14-H6	5.0	2.0	1.0
13-H8	14-H6	5.0	2.0	1.0
13-H2’	15-H8	6.0	2.0	1.0
…				
18-H2	19-H1’	5.0	2.0	1.0
18-H1’	19-H8	5.0	2.0	1.0
18-H2’	19-H8	3.8	2.0	0.7
18-H3’	19-H8	3.8	2.0	0.7
18-H8	19-H8	3.8	2.0	0.7
G19				
19-H1’	19-H8	5.0	2.0	1.0
19-H2’	19-H8	5.0	2.0	1.0
19-H2’	20-H1’	5.0	2.0	1.0
19-H2’	20-H2’	6.0	2.0	1.0
19-H8	20-H5	5.0	2.0	1.0
19-H1’	20-H6	5.0	3.0	1.0
19-H2’	20-H6	2.5	0.7	0.5
19-H8	20-H6	5.0	2.0	1.0
U20				
20-H2’	20-H5	5.0	2.0	1.0
20-H3’	20-H5	5.0	2.0	1.0
20-H1’	20-H6	3.8	2.0	0.7
20-H2’	20-H6	3.8	2.0	0.7
20-H3’	20-H6	3.8	2.0	0.7
20-H2’	21-H1’	5.0	2.0	1.0
20-H2’	21-H2’	6.0	2.0	1.0
20-H1’	21-H8	5.0	3.0	1.0
20-H2’	21-H8	2.5	0.7	0.5
20-H3’	21-H8	3.8	2.0	0.7
20-H6	21-H8	5.0	2.0	1.0
G21				
21-H1’	21-H8	3.8	2.0	0.7
21-H2’	21-H8	3.8	2.0	0.7
21-H2’	22-H1’	5.0	2.0	1.0
21-H1’	22-H8	6.0	3.0	1.0
21-H2’	22-H8	5.0	2.0	1.0
21-H2’	23-H8	3.8	2.0	1.0
A22				
22-H1’	22-H8	3.8	2.0	0.7
22-H2’	22-H8	5.0	2.0	2.0
22-H2’	23-H1’	5.0	2.0	1.0
22-H2	23-H1’	3.8	2.0	0.7
22-H1’	23-H8	5.0	2.0	1.0
22-H2’	23-H8	3.8	2.0	0.7
22-H8	23-H8	5.0	2.0	1.0
22-H1’	39-H2	6.0	2.0	1.0
22-H2	39-H2	4.0	2.0	1.0
A23				
23-H1’	23-H8	5.0	2.0	1.0
23-H2’	23-H8	5.0	2.0	1.0
U24				
24-H1’	24-H6	3.8	2.0	0.7
24-H2’	24-H6	5.0	3.0	1.0
24-H2’	25-H1’	5.0	2.0	1.0
24-H1’	25-H6	5.0	2.0	1.0
24-H2’	25-H6	3.8	2.0	0.7
24-H6	25-H6	5.0	2.0	1.0
…				
25-H1’	37-H2	3.8	2.0	0.7
25-H2’	37-H2	6.0	2.0	1.0
…				
35-H1’	36-H1’	6.0	2.0	1.0
35-H2’	36-H1’	5.0	2.0	1.0
35-H1’	36-H8	5.0	2.0	1.0
35-H2’	36-H8	2.5	0.7	0.5
35-H3’	36-H8	3.8	2.0	0.7
35-H6	36-H8	5.0	2.0	1.0
G36				
36-H1’	36-H8	5.0	2.0	1.0
36-H2’	36-H8	5.0	2.0	1.0
36-H1’	37-H8	3.8	2.0	0.7
36-H2’	37-H8	3.8	2.0	0.7
36-H8	37-H8	5.0	2.0	1.0
A37				
37-H1’	37-H8	5.0	2.0	1.0
37-H2’	37-H8	5.0	2.0	1.0

37-H2	38-H1'	3.8	2.0	0.7
37-H2'	38-H8	3.8	2.0	0.7
G38				
38-H1'	39-H1'	6.0	2.0	1.0
38-H1'	39-H8	5.0	3.0	1.0
38-H2'	39-H8	3.8	2.0	0.7
A39				
39-H1'	39-H2	6.0	2.0	1.0
39-H1'	39-H8	5.0	2.0	1.0
39-H2'	39-H8	5.0	2.0	2.0
39-H1'	40-H8	3.8	2.0	0.7
39-H2	40-H8	6.0	2.0	2.0
39-H2'	40-H8	3.8	2.0	0.7
39-H8	40-H8	6.0	3.0	1.0

NOE-derived distance restraints for exchangeable protons of D1kz in the absence of $MgCl_2$ (for all regions except the base pairs around κ) and in the presence of 10 mM $MgCl_2$ (for the base pairs next to κ):

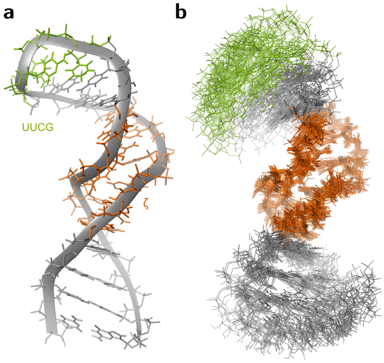
G1				
1-H1	2-H1	5.0	3.0	2.0
1-H1	2-H1'	5.0	3.0	2.0
1-H1	48-H41	5.0	3.0	2.0
1-H1	49-H1'	6.0	3.0	2.0
1-H1	49-H41	2.5	0.7	0.5
1-H1	49-H42	3.8	2.0	0.7
1-H1	49-H5	6.0	3.0	2.0
1-H1	49-H6	6.0	3.0	2.0
G2				
2-H1	3-H2	3.8	2.0	0.7
2-H1	3-H1'	6.0	3.0	2.0
2-H1	47-H3	3.8	2.0	0.7
2-H1	48-H1'	6.0	3.0	2.0
2-H1	48-H2'	6.0	3.0	2.0
2-H1	48-H41	2.5	0.7	0.5
2-H1	48-H42	3.8	2.0	0.7
2-H1	48-H5	5.0	3.0	2.0
2-H1	49-H1'	6.0	3.0	2.0
2-H1	49-H41	5.0	3.0	2.0
2-H1	49-H5	5.0	3.0	2.0
2-H1	49-H6	6.0	3.0	2.0
A3				
3-H2	46-H3	6.0	3.0	2.0
3-H2	47-H3	2.5	0.7	0.5
3-H61	47-H3	2.5	0.7	0.5
3-H62	47-H3	3.8	2.0	0.7
A4				
4-H2	45-H1	5.0	3.0	2.0
4-H62	45-H1	5.0	3.0	2.0
4-H2	46-H3	2.5	0.7	0.5
4-H61	46-H3	3.8	2.0	0.7
4-H62	46-H3	3.8	2.0	0.7
4-H1'	47-H3	6.0	3.0	1.0
4-H2	47-H3	3.8	2.0	0.7
U5				
5-H3	45-H1	2.5	0.7	0.5
5-H1'	46-H3	6.0	2.0	1.0
A6				
6-H1'	45-H1	6.0	3.0	2.0
G8				
8-H1	9-H1'	6.0	3.0	2.0
8-H1	9-H41	5.0	3.0	2.0
8-H1	9-H5	6.0	3.0	2.0
8-H1	9-H6	6.0	3.0	2.0
8-H1	40-H1	3.8	2.0	0.7
8-H1	41-H1'	5.0	3.0	2.0
8-H1	41-H41	3.8	2.0	0.7
8-H1	41-H42	3.8	2.0	0.7
8-H1	41-H5	5.0	3.0	2.0
8-H1	41-H6	6.0	3.0	2.0
8-H1	42-H5	5.0	3.0	2.0
8-H1	42-H6	6.0	3.0	2.0
C9				
9-H41	10-H3	3.8	2.0	0.7
9-H42	10-H3	5.0	3.0	2.0
9-H42	10-H5	5.0	3.0	2.0
9-H41	40-H1	2.5	0.7	0.5
9-H42	40-H1	3.8	2.0	0.7
9-H5	40-H1	5.0	3.0	2.0
9-H6	40-H1	6.0	3.0	2.0
U10				
10-H3	11-H1'	5.0	3.0	2.0
10-H3	11-H41	5.0	3.0	2.0
10-H3	11-H5	5.0	3.0	2.0
10-H3	11-H6	6.0	3.0	2.0
10-H3	22-H2	5.0	3.0	2.0
10-H3	23-H1'	6.0	3.0	2.0
10-H3	38-H1	3.8	2.0	0.7
10-H3	39-H2	2.5	0.7	0.5
10-H3	40-H1	3.8	2.0	0.7
10-H6	40-H1	6.0	3.0	2.0
C11				
11-H41	38-H1	3.8	2.0	0.7
11-H42	38-H1	5.0	3.0	2.0
11-H5	38-H1	5.0	3.0	2.0
11-H6	38-H1	6.0	3.0	2.0
A12				
12-H8	21-H1	6.0	3.0	2.0
A13				
13-H2	19-H1	3.8	2.0	0.7
13-H2	20-H3	2.5	0.7	0.5
13-H2	21-H1	6.0	3.0	2.0
C14				
14-H42	15-H1	5.0	3.0	2.0
14-H41	19-H1	2.5	0.7	0.5
14-H42	19-H1	3.8	2.0	0.7
14-H5	19-H1	5.0	3.0	2.0
14-H6	19-H1	6.0	3.0	2.0
14-H41	20-H3	3.8	2.0	0.7
14-H42	20-H3	3.8	2.0	0.7
G15				
15-H1	17-H8	5.0	3.0	2.0
15-H1	18-H8	6.0	2.0	2.0
15-H1'	19-H1	6.0	3.0	2.0
G19				
19-H1	20-H1'	6.0	3.0	2.0
19-H1	20-H3	3.8	2.0	0.7
19-H1	20-H6	6.0	3.0	2.0
U20				
20-H3	20-H1'	5.0	3.0	2.0
20-H3	21-H1	5.0	3.0	2.0
G21				
21-H1	23-H8	5.0	3.0	1.0
U24				
24-H3	25-H41	5.0	3.0	2.0
24-H3	25-H42	5.0	3.0	2.0
24-H3	25-H5	5.0	3.0	2.0
24-H3	36-H1	3.8	2.0	0.7
24-H3	37-H2	2.5	0.7	0.5
24-H3	38-H1	5.0	3.0	2.0
C25				
25-H41	35-H3	6.0	3.0	2.0
25-H41	36-H1	2.5	0.7	0.5
25-H42	36-H1	3.8	2.0	0.7
25-H5	36-H1	5.0	3.0	2.0
25-H6	36-H1	6.0	3.0	2.0
A26				
26-H2	35-H3	2.5	0.7	0.5
26-H8	36-H1	6.0	3.0	2.0
G27				
27-H1	28-H6	6.0	3.0	2.0
27-H1	34-H41	2.5	0.7	0.5
27-H1	34-H42	3.8	2.0	0.7
27-H1	34-H5	5.0	3.0	2.0
27-H1	34-H6	6.0	3.0	2.0
27-H1	35-H6	6.0	3.0	2.0
27-H8	35-H3	6.0	3.0	2.0
C28				
28-H41	29-H1'	6.0	3.0	2.0
28-H41	29-H5	6.0	3.0	2.0
28-H42	29-H5	5.0	3.0	2.0
28-H41	33-H1	2.5	0.7	0.5
28-H41	33-H22	5.0	3.0	2.0
28-H42	33-H1	3.8	2.0	0.7
28-H2'	33-H1	5.0	3.0	2.0
28-H5	33-H1	5.0	3.0	2.0
28-H6	33-H1	6.0	3.0	2.0
28-H2'	33-H21	5.0	3.0	2.0
U29				
29-H3	31-H1'	5.0	3.0	2.0

29-H3	31-H6	6.0	3.0	2.0
29-H1’	31-H41	6.0	3.0	2.0
29-H1’	31-H42	6.0	3.0	2.0
29-H2’	31-H41	3.8	2.0	0.7
29-H2’	31-H42	3.8	2.0	0.7
29-H5	31-H41	5.0	3.0	2.0
29-H5	31-H42	5.0	3.0	2.0
29-H6	31-H41	5.0	3.0	2.0
29-H1’	32-H1	3.8	2.0	0.7
29-H1’	32-H21	5.0	3.0	2.0
29-H2’	32-H1	3.8	2.0	0.7
29-H2’	32-H21	5.0	3.0	2.0
29-H3	32-H1	5.0	3.0	1.0
29-H1’	33-H1	5.0	3.0	2.0
29-H1’	33-H21	3.8	2.0	0.7
29-H1’	33-H22	3.8	2.0	0.7
29-H3	33-H1	5.0	3.0	2.0
29-H5	33-H1	6.0	3.0	2.0
U30				
30-H2’	32-H1	5.0	3.0	2.0
C31				
31-H41	32-H1	6.0	3.0	2.0
31-H42	32-H1	6.0	3.0	2.0
31-H1’	32-H1	3.8	2.0	0.7
31-H2’	32-H1	6.0	3.0	2.0
31-H5	32-H1	5.0	3.0	2.0
31-H5	32-H21	6.0	3.0	2.0
31-H6	32-H1	5.0	3.0	2.0
G32				
32-H1’	32-H1	6.0	3.0	2.0
32-H1	33-H1	5.0	3.0	2.0
32-H1	33-H8	6.0	3.0	2.0
32-H21	33-H1	5.0	3.0	2.0
32-H21	33-H21	6.0	3.0	2.0
G33				
33-H1	34-H1’	6.0	3.0	2.0
33-H21	34-H1’	3.8	2.0	0.7
33-H22	34-H1’	5.0	3.0	2.0
33-H1	34-H5	6.0	3.0	2.0
C34				
34-H41	35-H5	6.0	3.0	2.0
34-H42	35-H5	6.0	3.0	2.0
U35				
35-H3	36-H1	3.8	2.0	0.7
35-H3	36-H1’	6.0	3.0	2.0
35-H3	36-H8	6.0	3.0	2.0
G36				
36-H1	36-H1’	6.0	3.0	2.0
36-H1	37-H2	5.0	3.0	2.0
G38				
38-H1	39-H2	5.0	3.0	2.0
A39				
39-H2	40-H1	6.0	3.0	2.0
G40				
40-H1	41-H41	5.0	3.0	2.0
A44				
44-H2	45-H1	5.0	3.0	1.0
G45				
45-H1	46-H1’	6.0	3.0	2.0
45-H1	46-H3	5.0	3.0	2.0
45-H1	46-H5	6.0	3.0	2.0
U46				
46-H1’	46-H3	6.0	3.0	2.0
46-H3	47-H3	3.8	2.0	0.7
U47				
47-H1’	47-H3	6.0	3.0	2.0
47-H3	48-H41	5.0	3.0	2.0
47-H3	48-H42	5.0	3.0	2.0
47-H3	48-H5	5.0	3.0	2.0

Appendix 29. RDC restraints used in D1kz structure calculations.

first nucleus	second nucleus	RDC (Hz)	RDC error (Hz)
12-C2	12-H2	10.1	1.5
13-C8	13-H8	21.8	1.5
16-C8	16-H8	17.9	1.5
18-C2	18-H2	3.4	1.5
18-C8	18-H8	5.0	1.5
26-C2	26-H2	7.3	1.5
26-C8	26-H8	9.9	1.5
3-C2	3-H2	15.1	1.5
3-C8	3-H8	19.0	1.5
39-C2	39-H2	14.3	1.5
39-C8	39-H8	20.8	1.5
4-C8	4-H8	12.2	1.5
43-C2	43-H2	13.7	1.5
31-C1’	31-H1’	1.1	1.5
31-C5	31-H5	4.6	1.5
19-C8	19-H8	22.1	1.5
21-C8	21-H8	0.2	1.5
27-C8	27-H8	11.0	1.5
32-C1’	32-H1’	15.5	1.5
32-C8	32-H8	23.6	1.5
36-C8	36-H8	6.3	1.5
40-C8	40-H8	27.0	1.5
29-C1’	29-H1’	-19.6	1.5
30-C1’	30-H1’	18.5	1.5
30-C5	30-H5	12.4	1.5
30-C6	30-H6	2.6	1.5
47-C5	47-H5	17.8	1.5
41-C1’	41-H1’	-7.4	1.5
15-C8	15-H8	20.5	1.5
11-C1’	11-H1’	-2.4	1.5
13-C1’	13-H1’	-18.5	1.5
13-C2	13-H2	4.6	1.5
22-C8	22-H8	12.5	1.5
23-C8	23-H8	17.3	1.5
17-C8	17-H8	6.9	1.5

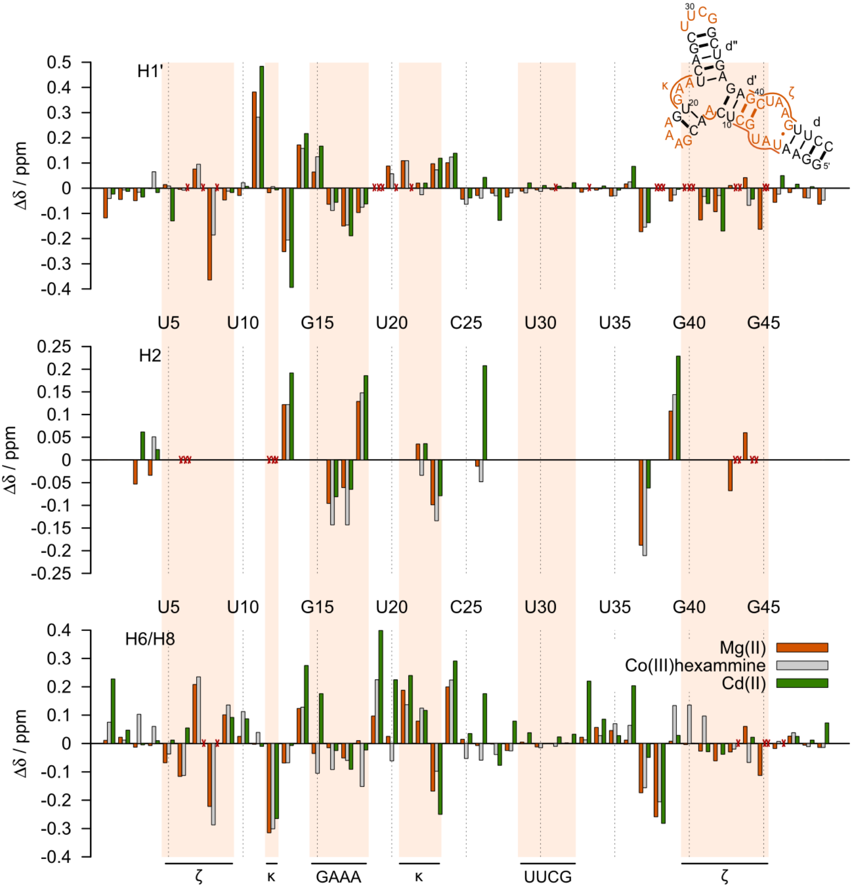
Appendix 30. D1-27 solution structure and structure calculation statistics. Orange: the tetraloop receptor (ζ), green: UUCG loop. a) Lowest energy structure of D1-27. b) Ensemble of the 20 lowest energy structures of D1-27 aligned over all heavy atoms of the tetraloop receptor.



RMSDs^{a}	
overall (1-27)	2.63 ± 1.01
helix d (1-5,23-27)	0.67 ± 0.32
ζ (5-9,18-23)	0.93 ± 0.39
helix d' and UUCG (8-19)	1.19 ± 0.56
distance restraints	
total	612
intranucleotide (i=j)	213
sequential (|i-j|=1)	293
medium (2≤|i-j|≤5)	29
long range (5<|i-j|)	77
dihedral restraints	173
H-bond restraints	47

[a] Root mean square deviations (RMSDs) for the ensembles of 20 lowest energy structures calculated in Xplor-NIH are given.

Appendix 31. Comparison of ^{1}H chemical shift changes in D1kz upon addition of 9 mM $MgCl_2$, 3 mM $Co(NH_3)_6Cl_3$ and 3 mM $Cd(ClO_4)_2$. To distinguish resonances that are not moving at all from values that are missing because assignment is not possible under all conditions the latter were crossed out red in the graph. κ and ζ as well as the GAAA and UUCG loops, are highlighted in orange in the graph and in the secondary structure scheme. (0.35-0.6 mM D1kz, 300 K, 60 mM K(I), pD 6.7-6.9, 10 µM EDTA, 700 MHz)



List of Figures

List of Tables

Bibliography

[1] Crick, F. H. (1958) On protein synthesis. *Symp. Soc. Exp. Biol. 12*, 138–163.

[2] Crick, F. H. (1970) Central dogma of molecular biology. *Nature 227*(5258), 561–563.

[3] Hüttenhofer, A., Schattner, P., and Polacek, N. (2005) Non-coding RNAs: hope or hype? *Trends Genet. 21*(5), 289–297.

[4] Hammond, S. M. (2005) Dicing and slicing: the core machinery of the RNA interference pathway. *FEBS Lett. 579*(26), 5822–5829.

[5] Cordaux, R., and Batzer, M. A. (2009) The impact of retrotransposons on human genome evolution. *Nat. Rev. Genet. 10*(10), 691–703.

[6] Kruger, K., Grabowski, P. J., Zaug, A. J., Sands, J., Gottschling, D. E., and Cech, T. R. (1982) Self-splicing RNA: autoexcision and autocyclization of the ribosomal RNA intervening sequence of Tetrahymena. *Cell 31*(1), 147–157.

[7] Guerrier-Takada, C., Gardiner, K., Marsh, T., Pace, N., and Altman, S. (1983) The RNA moiety of ribonuclease P is the catalytic subunit of the enzyme. *Cell 35*(3, Part 2), 849–857.

[8] Pyle, A. M. (1993) Ribozymes: a distinct class of metalloenzymes. *Science 261*(5122), 709–714.

[9] Donghi, D., and Schnabl, J. (2011) Multiple roles of metal ions in large ribozymes. *Met. Ions Life Sci. 9*, 197–234.

[10] Butcher, S. E., and Pyle, A. M. (2011) The molecular interactions that stabilize RNA tertiary structure: RNA motifs, patterns, and networks. *Acc. Chem. Res. 44*(12), 1302–1311.

[11] Milligan, J. F., Groebe, D. R., Witherell, G. W., and Uhlenbeck, O. C. (1987) Oligoribonucleotide synthesis using T7 RNA polymerase and synthetic DNA templates. *Nucleic Acids Res. 15*(21), 8783–8798.

[12] Milligan, J. F., and Uhlenbeck, O. C. (1989) Synthesis of small RNAs using T7 RNA polymerase. *Methods Enzymol. 180*, 51–62.

[13] Saenger, W. *Principles of nucleic acid structure*; Springer-Verlag, 1984.

[14] Harris, T. K., and Turner, G. J. (2002) Structural basis of perturbed pKa values of catalytic groups in enzyme active sites. *IUBMB life 53*(2), 85–98.

[15] Tang, C. L., Alexov, E., Pyle, A. M., and Honig, B. (2007) Calculation of pKas in RNA: on the structural origins and functional roles of protonated nucleotides. *J. Mol. Biol. 366*(5), 1475–1496.

[16] Arnott, S., Fuller, W., Hodgson, A., and Prutton, I. (1968) Molecular conformations and structure transitions of RNA complementary helices and their possible biological significance. *Nature 220*(5167), 561–564.

[17] Ghosh, A., and Bansal, M. (2003) A glossary of DNA structures from A to Z. *Acta Crystallogr. Sect. D Biol. Crystallogr. 59*(4), 620–626.

[18] Tamura, M., and Holbrook, S. R. (2002) Sequence and structural conservation in RNA ribose zippers. *J. Mol. Biol. 320*(3), 455–474.

[19] Leontis, N. B., and Westhof, E. (2001) Geometric nomenclature and classification of RNA base pairs. *RNA 7*(4), 499–512.

[20] Leontis, N. B., Stombaugh, J., and Westhof, E. (2002) The non-Watson-Crick base pairs and their associated isostericity matrices. *Nucleic Acids Res. 30*(16), 3497–3531.

[21] Nagaswamy, U., Larios-Sanz, M., Hury, J., Collins, S., Zhang, Z., Zhao, Q., and Fox, G. E. (2002) NCIR: a database of non-canonical interactions in known RNA structures. *Nucleic Acids Res. 30*(1), 395 –397.

[22] Stombaugh, J., Zirbel, C. L., Westhof, E., and Leontis, N. B. (2009) Frequency and isostericity of RNA base pairs. *Nucleic Acids Res. 37*(7), 2294–2312.

[23] Walberer, B. J., Cheng, A. C., and Frankel, A. D. (2003) Structural diversity and isomorphism of hydrogen-bonded base interactions in nucleic acids. *J. Mol. Biol. 327*(4), 767–780.

[24] Doherty, E. A., Batey, R. T., Masquida, B., and Doudna, J. A. (2001) A universal mode of helix packing in RNA. *Nat. Struct. Mol. Biol 8*(4), 339–343.

[25] Nissen, P., Ippolito, J. A., Ban, N., Moore, P. B., and Steitz, T. A. (2001) RNA tertiary interactions in the large ribosomal subunit: The A-minor motif. *Proc. Natl. Acad. Sci. U. S. A. 98*(9), 4899 –4903.

[26] Ban, N., Nissen, P., Hansen, J., Moore, P. B., and Steitz, T. A. (2000) The complete atomic structure of the large ribosomal subunit at 2.4 Å resolution. *Science 289*(5481), 905–920.

[27] Cate, J. H., Gooding, A. R., Podell, E., Zhou, K., Golden, B. L., Kundrot, C. E., Cech, T. R., and Doudna, J. A. (1996) Crystal structure of a group I ribozyme domain: principles of RNA packing. *Science 273*(5282), 1678–1685.

[28] Basu, S., Rambo, R. P., Strauss-Soukup, J., Cate, J. H. D., Ferré-D'Amaré, A. R., Strobel, S. A., and Doudna, J. A. (1998) A specific monovalent metal ion integral to the AA platform of the RNA tetraloop receptor. *Nat. Struct. Mol. Biol 5*(11), 986–992.

[29] Woese, C. R., Winker, S., and Gutell, R. R. (1990) Architecture of ribosomal RNA: constraints on the sequence of "tetra-loops". *Proc. Natl. Acad. Sci. 87*(21), 8467–8471.

[30] Wolters, J. (1992) The nature of preferred hairpin structures in 16S-like rRNA variable regions. *Nucleic Acids Res. 20*(8), 1843–1850.

[31] Costa, M., and Michel, F. (1997) Rules for RNA recognition of GNRA tetraloops deduced by in vitro selection: comparison with in vivo evolution. *EMBO J. 16*(11), 3289–3302.

[32] Ikawa, Y., Shiraishi, H., and Inoue, T. (1999) Trans-activation of the Tetrahymena group I intron ribozyme via a non-native RNA-RNA interaction. *Nucleic Acids Res. 27*(7), 1650–1655.

[33] Pley, H. W., Flaherty, K. M., and McKay, D. B. (1994) Model for an RNA tertiary interaction from the structure of an intermolecular complex between a GAAA tetraloop and an RNA helix. *Nature 372*(6501), 111–113.

[34] Abramovitz, D. L., and Pyle, A. M. (1997) Remarkable morphological variability of a common RNA folding motif: the GNRA Tetraloop-receptor interaction. *J. Mol. Biol. 266*(3), 493–506.

[35] Walter, A. E., Turner, D. H., Kim, J., Lyttle, M. H., Müller, P., Mathews, D. H., and Zuker, M. (1994) Coaxial stacking of helixes enhances binding of oligoribonucleotides and improves predictions of RNA folding. *Proc. Natl. Acad. Sci. U. S. A. 91*(20), 9218–9222.

[36] Walter, A. E., and Turner, D. H. (1994) Sequence dependence of stability for coaxial stacking of RNA helixes with Watson-Crick base paired interfaces. *Bio-*

chemistry 33(42), 12715–12719.

[37] Romaniuk, P. J., Hughes, D. W., Gregoire, R. J., Neilson, T., and Bell, R. A. (1978) Stabilizing effect of dangling bases on a short RNA double helix as determined by proton nuclear magnetic resonance spectroscopy. *J. Am. Chem. Soc. 100*(12), 3971–3972.

[38] Burkard, M. E., Kierzek, R., and Turner, D. H. (1999) Thermodynamics of unpaired terminal nucleotides on short RNA helixes correlates with stacking at helix termini in larger RNAs. *J. Mol. Biol. 290*(5), 967–982.

[39] Lilley, D. M. J. (2000) Structures of helical junctions in nucleic acids. *Q. Rev. Biophys. 33*(02), 109–159.

[40] Al-Hashimi, H. M., and Walter, N. G. (2008) RNA dynamics: it is about time. *Curr. Opin. Struct. Biol. 18*(3), 321–329.

[41] Bothe, J. R., Nikolova, E. N., Eichhorn, C. D., Chugh, J., Hansen, A. L., and Al-Hashimi, H. M. (2011) Characterizing RNA dynamics at atomic resolution using solution-state NMR spectroscopy. *Nat. Methods 8*(11), 919–931.

[42] McCammon, J. A., and Harvey, S. C. *Dynamics of proteins and nucleic acids*; Cambridge University Press: Cambridge, Cambridgeshire; New York, 1988.

[43] Duchardt, E., and Schwalbe, H. (2005) Residue specific ribose and nucleobase dynamics of the cUUCGg RNA tetraloop motif by MNMR 13C relaxation. *J. Biomol. NMR 32*(4), 295–308.

[44] Eldho, N. V., and Dayie, K. T. (2007) Internal bulge and tetraloop of the catalytic domain 5 of a group II intron ribozyme are flexible: Implications for catalysis. *J. Mol. Biol. 365*(4), 930–944.

[45] Ferner, J., Villa, A., Duchardt, E., Widjajakusuma, E., Wöhnert, J., Stock, G., and Schwalbe, H. (2008) NMR and MD studies of the temperature-dependent dynamics of RNA YNMG-tetraloops. *Nucleic Acids Res. 36*(6), 1928 –1940.

[46] Frank, A. T., Stelzer, A. C., Al-Hashimi, H. M., and Andricioaei, I. (2009) Constructing RNA dynamical ensembles by combining MD and motionally decoupled NMR RDCs: new insights into RNA dynamics and adaptive ligand recognition. *Nucleic Acids Res. 37*(11), 3670–3679.

[47] Stelzer, A. C., Frank, A. T., Kratz, J. D., Swanson, M. D., Gonzalez-Hernandez, M. J., Lee, J., Andricioaei, I., Markovitz, D. M., and Al-Hashimi, H. M. (2011) Discovery of selective bioactive small molecules by targeting an RNA dynamic ensemble. *Nat. Chem. Biol. 7*(8), 553–559.

[48] Latham, M. P., Brown, D. J., McCallum, S. A., and Pardi, A. (2005) NMR methods for studying the structure and dynamics of RNA. *ChemBioChem 6*(9), 1492–1505.

[49] Scott, L. G., and Hennig, M. (2008) RNA structure determination by NMR. *Methods Mol. Biol. 452*, 29–61.

[50] Kim, I., Lukavsky, P. J., and Puglisi, J. D. (2002) NMR study of 100 kDa HCV IRES RNA using segmental isotope labeling. *J. Am. Chem. Soc. 124*(32), 9338–9339.

[51] Varani, G., Aboul-Ela, F., and Allain, F. H. T. (1996) NMR investigation of RNA structure. *Prog. Nucl. Magn. Reson. Spectrosc. 29*(1-2), 51–127.

[52] Wijmenga, S. S., and Buuren, B. N. (1998) The use of NMR methods for conformational studies of nucleic acids. *Prog. Nucl. Magn. Reson. Spectrosc. 32*(4), 287–387.

[53] Fürtig, B., Richter, C., Wöhnert, J., and Schwalbe, H. (2003) NMR spectroscopy of RNA. *ChemBioChem 4*(10), 936–962.

[54] Flinders, J., and Dieckmann, T. (2006) NMR spectroscopy of ribonucleic acids. *Prog. Nucl. Magn. Reson. Spectrosc. 48*(2-3), 137–159.

[55] Tolbert, T. J., and Williamson, J. R. (1996) Preparation of specifically deuterated RNA for NMR Studies using a combination of chemical and enzymatic synthesis. *J. Am. Chem. Soc. 118*(34), 7929–7940.

[56] Otting, G., and Wüthrich, K. (1989) Extended heteronuclear editing of 2D 1H NMR spectra of isotope-labeled proteins, using the X(w1, w2) double half filter. *J. Magn. Reson. 85*(3), 586–594.

[57] Wüthrich, K. *NMR of proteins and nucleic acids*; Wiley, New York, 1986.

[58] Luy, B., and Marino, J. P. (2000) Direct evidence for Watson-Crick base pairs in a dynamic region of RNA structure. *J. Am. Chem. Soc. 122*(33), 8095–8096.

[59] Grzesiek, S., Cordier, F., Dingley, A. J., and Thomas L. James, V. D. Scalar couplings across hydrogen bonds. In *Nuclear Magnetic Resonance of Biological Macromolecules Part A*; Academic Press, 2002; Vol. Volume 338, pp 111–133.

[60] Lipsitz, R. S., and Tjandra, N. (2004) Residual dipolar couplings in NMR structure analysis. *Annu. Rev. Biophys. Biomol. Struct. 33*, 387–413.

[61] Hansen, M. R., Mueller, L., and Pardi, A. (1998) Tunable alignment of macromolecules by filamentous phage yields dipolar coupling interactions. *Nat. Struct. Biol. 5*(12), 1065–1074.

[62] Hansen, M. R., Hanson, P., and Pardi, A. (2000) Filamentous bacteriophage for aligning RNA, DNA, and proteins for measurement of nuclear magnetic resonance dipolar coupling interactions. *Methods Enzymol. 317*, 220–240.

[63] Woodson, S. A. (2005) Metal ions and RNA folding: a highly charged topic with a dynamic future. *Curr. Opin. Chem. Biol. 9*(2), 104–109.

[64] Sigel, R. K. O., and Pyle, A. M. (2007) Alternative roles for metal ions in enzyme catalysis and the implications for ribozyme chemistry. *Chem. Rev. 107*(1), 97–113.

[65] Lippert, B. (2008) Ligand-pKa shifts through metals: Potential relevance to ribozyme chemistry. *Chem. Biodiv. 5*(8), 1455–1474.

[66] Smith, M. D., Mehdizadeh, R., Olive, J. E., and Collins, R. A. (2008) The ionic environment determines ribozyme cleavage rate by modulation of nucleobase pKa. *RNA 14*(9), 1942 –1949.

[67] Erat, M. C., and Sigel, R. K. O. (2008) Divalent metal ions tune the self-splicing reaction of the yeast mitochondrial group II intron Sc.ai5gamma. *J. Biol. Inorg. Chem. 13*(6), 1025–1036.

[68] Steiner, M., Rueda, D., and Sigel, R. K. O. (2009) Ca^{2+} induces the formation of two distinct subpopulations of group II intron molecules. *Angew. Chem., Int. Ed. 48*(51), 9739–9742.

[69] Cayley, S., Lewis, B. A., Guttman, H. J., and Record Jr, M. (1991) Characterization of the cytoplasm of Escherichia coli K-12 as a function of external osmolarity: Implications for protein-DNA interactions in vivo. *J. Mol. Biol. 222*(2), 281–300.

[70] London, R. E. (1991) Methods for measurement of intracellular magnesium: NMR and fluorescence. *Annu. Rev. Physiol. 53*(1), 241–258.

[71] Froschauer, E. M., Kolisek, M., Dieterich, F., Schweigel, M., and Schweyen, R. J. (2004) Fluorescence measurements of free Mg^{2+} by use of mag-fura 2 in Salmonella enterica. *FEMS Microbiol. Lett. 237*(1), 49–55.

[72] Draper, D. E. (2008) RNA folding: Thermodynamic and molecular descriptions of the roles of ions. *Biophys. J. 95*(12), 5489–5495.

[73] Colmenarejo, G., and Tinoco, I. (1999) Structure and thermodynamics of metal binding in the P5 helix of a group I intron ribozyme. *J. Mol. Biol. 290*(1), 119–135.

[74] Tanaka, Y., and Taira, K. (2005) Detection of RNA nucleobase metalation by NMR spectroscopy. *Chem. Commun.* 2069–2070.

[75] Gordon, P. M., and Piccirilli, J. A. (2001) Metal ion coordination by the AGC triad in domain 5 contributes to group II intron catalysis. *Nat. Struct. Biol. 8*(10), 893–898.

[76] Shan, S.-o., Kravchuk, A. V., Piccirilli, J. A., and Herschlag, D. (2001) Defining the catalytic metal ion interactions in the Tetrahymena ribozyme reaction. *Biochemistry 40*(17), 5161–5171.

[77] Cowan, J. A. (1992) Transition metals as probes of metal cofactors in nucleic acid biochemistry. *Comments Inorg. Chem. 13*(5), 293 – 312.

[78] Cowan, J. A. (1993) Metallobiochemistry of RNA. Co(NH3)63+ as a probe for Mg2+(aq) binding sites. *J. Inorg. Biochem. 49*(3), 171–175.

[79] Kieft, J. S., and Tinoco, I. (1997) Solution structure of a metal-binding site in the major groove of RNA complexed with cobalt(III)hexammine. *Structure 5*(5), 713–721.

[80] Rüdisser, S., and Tinoco, I. (2000) Solution structure of cobalt(III)hexammine complexed to the GAAA tetraloop, and metal-ion binding to G·A mismatches. *J. Mol. Biol. 295*(5), 1211–1223.

[81] Gonzalez, R. L., and Tinoco, I. (2001) Identification and characterization of metal ion binding sites in RNA. *Methods Enzymol. 338*, 421–443.

[82] Hargittai, M. R., and Musier-Forsyth, K. (2000) Use of terbium as a probe of tRNA tertiary structure and folding. *RNA 6*(11), 1672–1680.

[83] Sigel, R. K. O., and Pyle, A. M. (2003) Lanthanide ions as probes for metal ions in the structure and catalytic mechanism of ribozymes. *Met. Ions Biol. Syst. 40*, 477–512.

[84] Feigon, J., Butcher, S. E., Finger, L. D., and Hud, N. V. (2001) Solution nuclear magnetic resonance probing of cation binding sites on nucleic acids. *Methods Enzymol. 338*, 400–420.

[85] Rulíšek, L., and Šponer, J. (2003) Outer-shell and inner-shell coordination of phosphate group to hydrated metal ions (Mg2+, Cu^{2+}, Zn^{2+}, Cd2+) in the presence and absence of nucleobase. The role of nonelectrostatic effects. *J. Phys. Chem. B 107*(8), 1913–1923.

[86] Draper, D. E., Grilley, D., and Soto, A. M. (2005) Ions and RNA folding. *Annu. Rev. Biophys. Biomol. Struct. 34*, 221–243.

[87] Draper, D. E., and Misra, V. K. (1998) RNA shows its metal. *Nat. Struct. Mol. Biol 5*(11), 927–930.

[88] Misra, V. K., and Draper, D. E. (2000) Mg^{2+} binding to tRNA revisited: the nonlinear Poisson-Boltzmann model. *J. Mol. Biol. 299*(3), 813–825.

[89] Draper, D. E. (2004) A guide to ions and RNA structure. *RNA 10*(3), 335–343.

[90] Klein, D. J., Moore, P. B., and Steitz, T. A. (2004) The contribution of metal ions to the structural stability of the large ribosomal subunit. *RNA 10*(9), 1366 –1379.

[91] Freisinger, E., and Sigel, R. K. O. (2007) From nucleotides to ribozymes–A comparison of their metal ion binding properties. *Coord. Chem. Rev. 251*(13-14), 1834–1851.

[92] Laing, L. G., Gluick, T. C., and Draper, D. E. (1994) Stabilization of RNA structure by Mg ions: Specific and non-specific effects. *J. Mol. Biol. 237*(5), 577–587.

[93] Misra, V. K., and Draper, D. E. (2001) A thermodynamic framework for Mg^{2+} binding to RNA. *Proc. Natl. Acad. Sci. U. S. A. 98*(22), 12456 –12461.

[94] Batey, R. T., and Doudna, J. A. (2002) Structural and energetic analysis of metal ions essential to SRP signal recognition domain assembly. *Biochemistry 41*(39), 11703–11710.

[95] Lee, T.-S., Giambasu, G. M., Sosa, C. P., Martick, M., Scott, W. G., and York, D. M. (2009) Threshold occupancy and specific cation binding modes in the hammerhead ribozyme active site are required for active conformation. *J. Mol. Biol. 388*(1), 195–206.

[96] Ennifar, E., Yusupov, M., Walter, P., Marquet, R., Ehresmann, B., Ehresmann, C., and Dumas, P. (1999) The crystal structure of the dimerization initiation site of genomic HIV-1 RNA reveals an extended duplex with two adenine bulges. *Structure 7*(11), 1439–1449.

[97] Chin, K., Sharp, K. A., Honig, B., and Pyle, A. M. (1999) Calculating the electrostatic properties of RNA provides new insights into molecular interactions and function. *Nat. Struct. Biol. 6*(11), 1055–1061.

[98] Lavery, R., and Pullman, B. (1981) The molecular electrostatic potential and steric accessibility of A-DNA. *Nucleic Acids Res. 9*(18), 4677–4688.

[99] Robinson, H., Gao, Y. G., Sanishvili, R., Joachimiak, A., and Wang, A. H. (2000) Hexahydrated magnesium ions bind in the deep major groove and at the outer mouth of A-form nucleic acid duplexes. *Nucleic Acids Res. 28*(8), 1760–1766.

[100] Dann III, C. E., Wakeman, C. A., Sieling, C. L., Baker, S. C., Irnov, I., and Winkler, W. C. (2007) Structure and mechanism of a metal-sensing regulatory RNA. *Cell 130*(5), 878–892.

[101] Batey, R. T., Rambo, R. P., Lucast, L., Rha, B., and Doudna, J. A. (2000) Crystal structure of the ribonucleoprotein core of the signal recognition particle. *Science 287*(5456), 1232 –1239.

[102] Cate, J. H., Hanna, R. L., and Doudna, J. A. (1997) A magnesium ion core at the heart of a ribozyme domain. *Nat. Struct. Biol. 4*(7), 553–558.

[103] Stahley, M. R., Adams, P. L., Wang, J., and Strobel, S. A. (2007) Structural metals in the group I intron: a ribozyme with a multiple metal ion core. *J. Mol. Biol. 372*(1), 89–102.

[104] Ramesh, A., Wakeman, C. A., and Winkler, W. C. (2011) Insights into metalloregulation by M-box riboswitch RNAs via structural analysis of manganese-bound complexes. *J. Mol. Biol. 407*(4), 556–570.

[105] Toor, N., Rajashankar, K., Keating, K. S., and Pyle, A. M. (2008) Structural basis for exon recognition by a group II intron. *Nat. Struct. Mol. Biol 15*(11), 1221–1222.

[106] Marcia, M., and Pyle, A. M. (2012) Visualizing Group II Intron Catalysis through the Stages of Splicing. *Cell 151*(3), 497–507.

[107] Dayie, K. T., and Padgett, R. A. (2008) A glimpse into the active site of a group II intron and maybe the spliceosome, too. *RNA 14*(9), 1697–1703.

[108] Correll, C. C., Freeborn, B., Moore, P. B., and Steitz, T. A. (1997) Metals, motifs, and recognition in the crystal structure of a 5S rRNA domain. *Cell 91*(5), 705–712.

[109] Bukhman, Y. V., and Draper, D. E. (1997) Affinities and selectivities of divalent cation binding sites within an RNA tertiary structure. *J. Mol. Biol. 273*(5), 1020–1031.

[110] Horton, T. E., Clardy, D. R., and DeRose, V. J. (1998) Electron paramagnetic resonance spectroscopic measurement of Mn2+ binding affinities to the hammerhead ribozyme and correlation with cleavage activity. *Biochemistry 37*(51), 18094–18101.

[111] Das, R., Travers, K. J., Bai, Y., and Herschlag, D. (2005) Determining the Mg^{2+} stoichiometry for folding an RNA metal ion core. *J. Am. Chem. Soc. 127*(23), 8272–8273.

[112] Leipply, D., and Draper, D. E. (2011) Evidence for a thermodynamically distinct

Mg^{2+} ion associated with formation of an RNA tertiary structure. *J. Am. Chem. Soc. 133*(34), 13397–13405.

[113] Gonzalez, R. L., and Tinoco, I. (1999) Solution structure and thermodynamics of a divalent metal ion binding site in an RNA pseudoknot. *J. Mol. Biol. 289*(5), 1267–1282.

[114] Erat, M. C., and Sigel, R. K. O. (2007) Determination of the intrinsic affinities of multiple site-specific Mg^{2+} ions coordinated to domain 6 of a group II intron ribozyme. *Inorg. Chem. 46*(26), 11224–11234.

[115] Spano, M. N., and Walter, N. G. (2011) Solution structure of an alternate conformation of helix27 from Escherichia coli16S rRNA. *Biopolymers 95*(10), 653–668.

[116] Korth, M. M. T., and Sigel, R. K. O. (2012) Unusually high-affinity Mg^{2+} binding at the AU-rich sequence within the antiterminator hairpin of a Mg^{2+} riboswitch. *Chem. Biodiv. 9*(9), 2035–2049.

[117] Misra, V. K., and Draper, D. E. (2002) The linkage between magnesium binding and RNA folding. *J. Mol. Biol. 317*(4), 507–521.

[118] Conn, G. L., Gittis, A. G., Lattman, E. E., Misra, V. K., and Draper, D. E. (2002) A compact RNA tertiary structure contains a buried backbone-K+ complex. *J. Mol. Biol. 318*(4), 963–973.

[119] Collie, G. W., Haider, S. M., Neidle, S., and Parkinson, G. N. (2010) A crystallographic and modelling study of a human telomeric RNA (TERRA) quadruplex. *Nucleic Acids Res. 38*(16), 5569–5580.

[120] Fedor, M. J. (2002) The role of metal ions in RNA catalysis. *Curr. Opin. Struct. Biol. 12*(3), 289–295.

[121] Fedor, M. J. (2009) Comparative enzymology and structural biology of RNA self-cleavage. *Ann. Rev. Biophys. 38*(1), 271–299.

[122] Lönnberg, T. (2011) Understanding catalysis of phosphate-transfer reactions by the large ribozymes. *Chem.–Eur. J. 17*(26), 7140–7153.

[123] Steitz, T. A., and Steitz, J. A. (1993) A general two-metal-ion mechanism for catalytic RNA. *Proc. Natl. Acad. Sci. U. S. A. 90*(14), 6498–6502.

[124] Stahley, M. R., and Strobel, S. A. (2005) Structural evidence for a two-metal-ion mechanism of group I intron splicing. *Science 309*(5740), 1587 –1590.

[125] Michel, F., Costa, M., and Westhof, E. (2009) The ribozyme core of group II introns: a structure in want of partners. *Trends Biochem. Sci. 34*(4), 189–199.

[126] Curtis, E. A., and Bartel, D. P. (2001) The hammerhead cleavage reaction in monovalent cations. *RNA 7*(04), 546–552.

[127] Roychowdhury-Saha, M., and Burke, D. H. (2006) Extraordinary rates of transition metal ion-mediated ribozyme catalysis. *RNA 12*(10), 1846–1852.

[128] Boots, J. L., Canny, M. D., Azimi, E., and Pardi, A. (2008) Metal ion specificities for folding and cleavage activity in the Schistosoma hammerhead ribozyme. *RNA 14*(10), 2212 –2222.

[129] Klawuhn, K., Jansen, J. A., Souchek, J., Soukup, G. A., and Soukup, J. K. (2010) Analysis of metal ion dependence in glmS ribozyme self-cleavage and coenzyme binding. *ChemBioChem 11*(18), 2567–2571.

[130] Schnabl, J., and Sigel, R. K. O. (2010) Controlling ribozyme activity by metal ions. *Curr. Opin. Chem. Biol. 14*(2), 269–275.

[131] DeRose, V. J. Characterization of nucleic acid-metal ion binding by spectroscopic techniques. In *Nucleic Acid-Metal Ion Interactions*; Royal Society of Chemistry, Cambridge, UK, 2009; pp 154–179.

[132] Johannsen, S., Korth, M. M. T., Schnabl, J., and Sigel, R. K. O. (2009) Exploring metal ion coordination to nucleic acids by NMR. *Chimia 63*(3), 146–152.

[133] Johannsen, S., Megger, N., Böhme, D., Sigel, R. K. O., and Müller, J. (2010) Solution structure of a DNA double helix with consecutive metal-mediated base pairs. *Nature Chem. 2*(3), 229–234.

[134] Donghi, D., and Sigel, R. K. O. (2012) Metal ion-RNA interactions studied via multinuclear NMR. *Methods Mol. Biol. 848*, 253–273.

[135] Reid, S. S., and Cowan, J. A. (1990) Biostructural chemistry of magnesium ion: characterization of the weak binding sites on tRNAPhe(yeast). Implications for conformational change and activity. *Biochemistry 29*(25), 6025–6032.

[136] Cowan, J. A. (1991) Coordination chemistry of magnesium ions and 5S rRNA (Escherichia coli): binding parameters, ligand symmetry, and implications for activity. *J. Am. Chem. Soc. 113*(2), 675–676.

[137] Cowan, J. A., Huang, H.-W., and Hsu, L. Y. (1993) Sequence selective coordination of Mg2+(aq) to DNA. *J. Inorg. Biochem. 52*(2), 121–129.

[138] Wright, L. A., and Lerner, L. E. (1994) Magnesium-DNA interactions from interpretation of 25Mg-NMR relaxation rates: field and coion dependence. *Biopolymers 34*(6), 691–700.

[139] Berggren, E., Nordenskiöld, L., and Braunlin, W. H. (1992) Interpretation of 25Mg spin relaxation in Mg-DNA solutions: temperature variation and chemical exchange effects. *Biopolymers 32*(10), 1339–1350.

[140] Lian, L.-Y., and Roberts, G. C. K. Effects of chemical exchange on NMR spectra. In *NMR of macromolecules: a practical approach*; IRL Press at Oxford University Press: Oxford ;;New York, 1993; pp 153–182.

[141] Gutowsky, H. S., and Holm, C. H. (1956) Rate processes and nuclear magnetic resonance spectra. II. Hindered internal rotation of amides. *J. Chem. Phys. 25*(6), 1228.

[142] Pecoraro, V. L., Hermes, J. D., and Cleland, W. W. (1984) Stability constants of Mg^{2+} and Cd^{2+} complexes of adenine nucleotides and thionucleotides and rate constants for formation and dissociation of MgATP and MgADP. *Biochemistry 23*(22), 5262–5271.

[143] Gill, M. L., Strobel, S. A., and Loria, J. P. (2005) 205Tl NMR methods for the characterization of monovalent cation binding to nucleic acids. *J. Am. Chem. Soc. 127*(47), 16723–16732.

[144] Bryant, R. G. (1983) The NMR time scale. *J. Chem. Educ. 60*(11), 933.

[145] Erat, M. C., Coles, J., Finazzo, C., Knobloch, B., and Sigel, R. K. (2012) Accurate analysis of Mg^{2+} binding to RNA: From classical methods to a novel iterative calculation procedure. *Coord. Chem. Rev. 256*(1–2), 279–288.

[146] Wang, G., Gaffney, B. L., and Jones, R. A. (2004) Differential binding of Mg^{2+}, Zn^{2+}, and Cd^{2+} at two sites in a hammerhead ribozyme motif, determined by 15N NMR. *J. Am. Chem. Soc. 126*(29), 8908–8909.

[147] Sychrovský, V., Šponer, J., and Hobza, P. (2004) Theoretical calculation of the NMR spin-spin coupling constants and the NMR shifts allow distinguishability between the specific direct and the water-mediated binding of a divalent metal cation to guanine. *J. Am. Chem. Soc. 126*(2), 663–672.

[148] Li, H., Cukier, R. I., and Bu, Y. (2008) Remarkable metal counterion effect on the internucleotide J-couplings and chemical shifts of the N-H···N hydrogen bonds in the W-C base pairs. *J. Phys. Chem. B 112*(30), 9174–9181.

[149] Campbell, D. O., Bouchard, P., Desjardins, G., and Legault, P. (2006) NMR structure of Varkud Satellite ribozyme stem-loop V in the presence of magnesium ions and localization of metal-binding sites. *Biochemistry 45*(35), 10591–10605.

[150] Davis, J. H., Foster, T. R., Tonelli, M., and Butcher, S. E. (2007) Role of metal ions in the tetraloop-receptor complex as analyzed by NMR. *RNA 13*(1), 76–86.

[151] Noeske, J., Schwalbe, H., and Wöhnert, J. (2007) Metal-ion binding and metal-ion induced folding of the adenine-sensing riboswitch aptamer domain. *Nucleic Acids Res. 35*(15), 5262 –5273.

[152] Butcher, S. E., Allain, F. H.-T., and Feigon, J. (2000) Determination of metal ion binding sites within the hairpin ribozyme domains by NMR. *Biochemistry 39*(9), 2174–2182.

[153] Cech, T. R., Zaug, A. J., and Grabowski, P. J. (1981) In vitro splicing of the ribosomal RNA precursor of Tetrahymena: involvement of a guanosine nucleotide in the excision of the intervening sequence. *Cell 27*(3 Pt 2), 487–496.

[154] Gilbert, W. (1986) Origin of life: The RNA world. *Nature 319*(6055), 618–618.

[155] Kikovska, E., Svärd, S. G., and Kirsebom, L. A. (2007) Eukaryotic RNase P RNA mediates cleavage in the absence of protein. *Proc. Natl. Acad. Sci. 104*(7), 2062–2067.

[156] Collins, C. A., and Guthrie, C. (2000) The question remains: is the spliceosome a ribozyme? *Nat. Struct. Biol. 7*(10), 850–854.

[157] Nissen, P., Hansen, J., Ban, N., Moore, P. B., and Steitz, T. A. (2000) The structural basis of ribosome activity in peptide bond synthesis. *Science 289*(5481), 920–930.

[158] Hansen, J. L., Schmeing, T. M., Moore, P. B., and Steitz, T. A. (2002) Structural insights into peptide bond formation. *Proc. Natl. Acad. Sci. 99*(18), 11670–11675.

[159] Shukla, G. C., and Padgett, R. A. (2002) A catalytically active group II intron domain 5 can function in the U12-dependent spliceosome. *Mol. Cell 9*(5), 1145–1150.

[160] Valadkhan, S. (2007) The spliceosome: a ribozyme at heart? *Biol. Chem. 388*(7), 693–697.

[161] Bashan, A., and Yonath, A. (2008) Correlating ribosome function with high-resolution structures. *Trends Microbiol. 16*(7), 326–335.

[162] Butcher, S. E. (2009) The spliceosome as ribozyme hypothesis takes a second step. *Proc. Natl. Acad. Sci. 106*(30), 12211 –12212.

[163] Lee, C., Jaladat, Y., Mohammadi, A., Sharifi, A., Geisler, S., and Valadkhan, S. (2010) Metal binding and substrate positioning by evolutionarily invariant U6 sequences in catalytically active protein-free snRNAs. *RNA 16*(11), 2226–2238.

[164] Sontheimer, E. J., Gordon, P. M., and Piccirilli, J. A. (1999) Metal ion catalysis during group II intron self-splicing: parallels with the spliceosome. *Genes Dev. 13*(13), 1729–1741.

[165] Gordon, P. M., Sontheimer, E. J., and Piccirilli, J. A. (2000) Metal ion catalysis during the exon-ligation step of nuclear pre-mRNA splicing: extending the parallels between the spliceosome and group II introns. *RNA 6*(2), 199–205.

[166] Keating, K. S., Toor, N., Perlman, P. S., and Pyle, A. M. (2010) A structural analysis of the group II intron active site and implications for the spliceosome. *RNA 16*(1), 1–9.

[167] Joyce, G. F. (2004) Directed evolution of nucleic acid enzymes. *Annu. Rev. Biochem. 73*(1), 791–836.

[168] Ellington, A. D., and Szostak, J. W. (1990) In vitro selection of RNA molecules that bind specific ligands. *Nature 346*(6287), 818–822.

[169] Tarasow, T. M., Tarasow, S. L., and Eaton, B. E. (1997) RNA-catalysed carbon-carbon bond formation. *Nature 389*(6646), 54–57.

[170] Murakami, H., Saito, H., and Suga, H. (2003) A versatile tRNA aminoacylation catalyst based on RNA. *Chem. Biol. 10*(7), 655–662.

[171] Chumachenko, N. V., Novikov, Y., and Yarus, M. (2009) Rapid and simple

ribozymic aminoacylation using 3 conserved nucleotides. *J. Am. Chem. Soc. 131*(14), 5257–5263.

[172] Turk, R. M., Chumachenko, N. V., and Yarus, M. (2010) Multiple translational products from a five-nucleotide ribozyme. *Proc. Natl. Acad. Sci. U. S. A. 107*(10), 4585–4589.

[173] Ferré-D'Amaré, A. R., and Scott, W. G. (2010) Small self-cleaving ribozymes. *Cold Spring Harbor Perspectives in Biology 2*(10), a003574.

[174] Fedor, M. J. (2000) Structure and function of the hairpin ribozyme. *J. Mol. Biol. 297*(2), 269–291.

[175] Blount, K. F., and Uhlenbeck, O. C. (2005) The structure-function dilemma of the hammerhead ribozyme. *Annu. Rev. Biophys. Biomol. Struct. 34*(1), 415–440.

[176] Shih, I.-h., and Been, M. D. (2002) Catalytic strategies of the Hepatitis Delta Virus ribozymes. *Annu. Rev. Biochem. 71*(1), 887–917.

[177] Lilley, D. M. J. (2004) The Varkud satellite ribozyme. *RNA 10*(2), 151–158.

[178] Cochrane, J. C., and Strobel, S. A. (2008) Catalytic strategies of self-cleaving ribozymes. *Acc. Chem. Res. 41*(8), 1027–1035.

[179] Nakano, S.-i., Chadalavada, D. M., and Bevilacqua, P. C. (2000) General acid-base catalysis in the mechanism of a Hepatitis Delta Virus ribozyme. *Science 287*(5457), 1493 –1497.

[180] Bevilacqua, P. C., Brown, T. S., Nakano, S.-i., and Yajima, R. (2004) Catalytic roles for proton transfer and protonation in ribozymes. *Biopolymers 73*(1), 90–109.

[181] Bevilacqua, P. C., and Yajima, R. (2006) Nucleobase catalysis in ribozyme mechanism. *Curr. Opin. Chem. Biol. 10*(5), 455–464.

[182] Han, J., and Burke, J. M. (2005) Model for general acid-base catalysis by the hammerhead ribozyme: pH-activity relationships of G8 and G12 variants at the putative active site. *Biochemistry 44*(21), 7864–7870.

[183] Wilson, T. J., Li, N.-S., Lu, J., Frederiksen, J. K., Piccirilli, J. A., and Lilley, D. M. J. (2010) Nucleobase-mediated general acid-base catalysis in the Varkud satellite ribozyme. *Proc. Natl. Acad. Sci. 107*(26), 11751–11756.

[184] Murray, J. B., Seyhan, A. A., Walter, N. G., Burke, J. M., and Scott, W. G. (1998) The hammerhead, hairpin and VS ribozymes are catalytically proficient in monovalent cations alone. *Chem. Biol. 5*(10), 587–595.

[185] O'Rear, J. L., Wang, S., Feig, A. L., Beigelman, L., Uhlenbeck, O. C., and Herschlag, D. (2001) Comparison of the hammerhead cleavage reactions stimulated by monovalent and divalent cations. *RNA 7*(4), 537–545.

[186] Roth, A., Nahvi, A., Lee, M., Jona, I., and Breaker, R. R. (2006) Characteristics of the glmS ribozyme suggest only structural roles for divalent metal ions. *RNA 12*(4), 607–619.

[187] Perrotta, A. T., and Been, M. D. (2006) HDV ribozyme activity in monovalent cations. *Biochemistry 45*(38), 11357–11365.

[188] Ke, A., Ding, F., Batchelor, J. D., and Doudna, J. A. (2007) Structural roles of monovalent cations in the HDV ribozyme. *Structure 15*(3), 281–287.

[189] Serganov, A., and Patel, D. J. (2007) Ribozymes, riboswitches and beyond: regulation of gene expression without proteins. *Nat. Rev. Genet. 8*(10), 776–790.

[190] Winkler, W. C., Nahvi, A., Roth, A., Collins, J. A., and Breaker, R. R. (2004) Control of gene expression by a natural metabolite-responsive ribozyme. *Nature 428*(6980), 281–286.

[191] Cochrane, J. C., Lipchock, S. V., and Strobel, S. A. (2007) Structural investigation of the GlmS ribozyme bound to Its catalytic cofactor. *Chem. Biol. 14*(1), 97–105.

[192] Salehi-Ashtiani, K., Lupták, A., Litovchick, A., and Szostak, J. W. (2006) A genomewide search for ribozymes reveals an HDV-like sequence in the human CPEB3 gene. *Science 313*(5794), 1788–1792.

[193] Webb, C.-H. T., Riccitelli, N. J., Ruminski, D. J., and Lupták, A. (2009) Widespread occurrence of self-cleaving ribozymes. *Science 326*(5955), 953–953.

[194] Teixeira, A., Tahiri-Alaoui, A., West, S., Thomas, B., Ramadass, A., Martianov, I., Dye, M., James, W., Proudfoot, N. J., and Akoulitchev, A. (2004) Autocatalytic RNA cleavage in the human β-globin pre-mRNA promotes transcription termination. *Nature 432*(7016), 526–530.

[195] Woodson, S. A. (2005) Structure and assembly of group I introns. *Curr. Opin. Struct. Biol. 15*(3), 324–330.

[196] Nielsen, H., and Johansen, S. D. (2009) Group I introns: Moving in new directions. *RNA Biol. 6*(4), 375–383.

[197] Lehmann, K., and Schmidt, U. (2003) Group II introns: structure and catalytic versatility of large natural ribozymes. *Crit. Rev. Biochem. Mol. Biol. 38*(3), 249–303.

[198] Pyle, A. M. (2010) The tertiary structure of group II introns: implications for biological function and evolution. *Crit. Rev. Biochem. Mol. Biol. 45*(3), 215–232.

[199] McClain, W. H., Lai, L. B., and Gopalan, V. (2010) Trials, travails and triumphs: An account of RNA catalysis in RNase P. *J. Mol. Biol. 397*(3), 627–646.

[200] Smith, D., and Pace, N. R. (1993) Multiple magnesium ions in the ribonuclease P reaction mechanism. *Biochemistry 32*(20), 5273–5281.

[201] Cassano, A. G., Anderson, V. E., and Harris, M. E. (2004) Analysis of solvent nucleophile isotope effects: evidence for concerted mechanisms and nucleophilic activation by metal coordination in nonenzymatic and ribozyme-catalyzed phosphodiester hydrolysis. *Biochemistry 43*(32), 10547–10559.

[202] Talini, G., Gallori, E., and Maurel, M.-C. (2009) Natural and unnatural ribozymes: Back to the primordial RNA world. *Res. Microbiol. 160*(7), 457–465.

[203] Fedor, M. J., and Williamson, J. R. (2005) The catalytic diversity of RNAs. *Nat. Rev. Mol. Cell Biol. 6*(5), 399–412.

[204] Strobel, S. A., and Cochrane, J. C. (2007) RNA catalysis: ribozymes, ribosomes, and riboswitches. *Curr. Opin. Chem. Biol. 11*(6), 636–643.

[205] Kurz, J. C., and Fierke, C. A. (2000) Ribonuclease P: a ribonucleoprotein enzyme. *Curr. Opin. Chem. Biol. 4*(5), 553–558.

[206] Pyle, A. (2005) Capping by branching: A new ribozyme makes tiny lariats. *Science 309*(5740), 1530–1531.

[207] Nielsen, H., Westhof, E., and Johansen, S. (2005) Molecular biology: An mRNA is capped by a 2',5' lariat catalyzed by a group I-like ribozyme. *Science 309*(5740), 1584–1587.

[208] Gordon, P. M., Fong, R., and Piccirilli, J. A. (2007) A second divalent metal ion in the group II intron reaction center. *Chem. Biol. 14*(6), 607–612.

[209] Roitzsch, M., Fedorova, O., and Pyle, A. M. (2010) The 2'-OH group at the group II intron terminus acts as a proton shuttle. *Nat. Chem. Biol. 6*(3), 218–224.

[210] Ferat, J.-L., and Michel, F. (1993) Group II self-splicing introns in bacteria. *Nature 364*(6435), 358–361.

[211] Martínez-Abarca, F., and Toro, N. (2000) Group II introns in the bacterial world. *Mol. Microbiol. 38*(5), 917–926.

[212] Toro, N., Molina-Sánchez, M. D., and Fernández-López, M. (2002) Identification and characterization of bacterial class E group II introns. *Gene 299*(1-2), 245–250.

[213] Michel, F., Umesono, K., and Ozeki, H. (1989) Comparative and functional anatomy of group II catalytic introns — a review. *Gene 82*(1), 5–30.

[214] Dellaporta, S. L., Xu, A., Sagasser, S., Jakob, W., Moreno, M. A., Buss, L. W., and Schierwater, B. (2006) Mitochondrial genome of Trichoplax adhaerens supports Placozoa as the basal lower metazoan phylum. *Proc. Natl. Acad. Sci. 103*(23), 8751–8756.

[215] Vallès, Y., Halanych, K. M., and Boore, J. L. (2008) Group II introns break new boundaries: presence in a bilaterian's genome. *PloS One 3*(1), e1488.

[216] Burger, G., Yan, Y., Javadi, P., and Lang, B. F. (2009) Group I-intron trans-splicing and mRNA editing in the mitochondria of placozoan animals. *Trends Genet. 25*(9), 381–386.

[217] Chin, K., and Pyle, A. M. (1995) Branch-point attack in group II introns is a highly reversible transesterification, providing a potential proofreading mechanism for 5'-splice site selection. *RNA 1*(4), 391–406.

[218] Lambowitz, A. M., and Zimmerly, S. (2004) Mobile group II introns. *Annu. Rev. Genet. 38*, 1–35.

[219] Koonin, E. V. (2006) The origin of introns and their role in eukaryogenesis: a compromise solution to the introns-early versus introns-late debate? *Biol. Direct 1*, 22.

[220] Robart, A. R., and Zimmerly, S. (2005) Group II intron retroelements: function and diversity. *Cytogenet. Genome Res. 110*(1-4), 589–597.

[221] Beauregard, A., Curcio, M. J., and Belfort, M. (2008) The take and give between retrotransposable elements and their hosts. *Annu. Rev. Genet. 42*, 587–617.

[222] Jones, J. P., Kierlin, M. N., Coon, R. G., Perutka, J., Lambowitz, A. M., and Sullenger, B. A. (2005) Retargeting mobile group II introns to repair mutant genes. *Mol. Ther. 11*(5), 687–694.

[223] Yao, J., and Lambowitz, A. M. (2007) Gene targeting in gram-negative bacteria by use of a mobile group II intron ("targetron") expressed from a broad-host-range vector. *Appl. Environ. Microbiol. 73*(8), 2735–2743.

[224] Nazari, R., and Joshi, S. (2008) Exploring the potential of group II introns to inactivate human immunodeficiency virus type 1. *J. Gen. Virol. 89*(Pt 10), 2605–2610.

[225] Mastroianni, M., Watanabe, K., White, T., Zhuang, F., Vernon, J., Matsuura, M., Wallingford, J., and Lambowitz, A. (2008) Group II intron-based gene targeting reactions in eukaryotes. *PloS One 3*(9), e3121.

[226] Zhuang, F., Karberg, M., Perutka, J., and Lambowitz, A. M. (2009) EcI5, a group IIB intron with high retrohoming frequency: DNA target site recognition and use in gene targeting. *RNA 15*(3), 432–449.

[227] Boeke, J. D. (2003) The unusual phylogenetic distribution of retrotransposons: A hypothesis. *Genome Res. 13*(9), 1975 –1983.

[228] Sigel, R. K. O. (2005) Group II Intron ribozymes and metal ions - a delicate relationship. *Eur. J. Inorg. Chem. 2005*(12), 2281–2292.

[229] Solem, A., Zingler, N., Pyle, A. M., and Pook-Than, J. Group II introns and their protein collaborators. In *Non-Protein Coding RNAs*; Walter, N. G., Woodson, S. A., and Batey, R. T., Eds.; Springer Berlin Heidelberg, 2009; Vol. 13, pp 167–182.

[230] Huang, H.-R., Rowe, C. E., Mohr, S., Jiang, Y., Lambowitz, A. M., and Perlman, P. S. (2005) The splicing of yeast mitochondrial group I and group II introns requires a DEAD-box protein with RNA chaperone function. *Proc. Natl. Acad. Sci. U. S. A. 102*(1), 163–168.

[231] Mohr, S., Matsuura, M., Perlman, P. S., and Lambowitz, A. M. (2006) A DEAD-

box protein alone promotes group II intron splicing and reverse splicing by acting as an RNA chaperone. *Proc. Natl. Acad. Sci. U. S. A. 103*(10), 3569–3574.

[232] Matsuura, M., Noah, J. W., and Lambowitz, A. M. (2001) Mechanism of maturase-promoted group II intron splicing. *EMBO J. 20*(24), 7259–7270.

[233] Toor, N., Hausner, G., and Zimmerly, S. (2001) Coevolution of group II intron RNA structures with their intron-encoded reverse transcriptases. *RNA 7*(8), 1142–1152.

[234] Simon, D. M., Clarke, N. A. C., McNeil, B. A., Johnson, I., Pantuso, D., Dai, L., Chai, D., and Zimmerly, S. (2008) Group II introns in eubacteria and archaea: ORF-less introns and new varieties. *RNA 14*(9), 1704–1713.

[235] Candales, M. A., Duong, A., Hood, K. S., Li, T., Neufeld, R. A. E., Sun, R., McNeil, B. A., Wu, L., Jarding, A. M., and Zimmerly, S. (2012) Database for bacterial group II introns. *Nucleic Acids Res. 40*(D1), D187–D190.

[236] Toor, N., Keating, K. S., Taylor, S. D., and Pyle, A. M. (2008) Crystal structure of a self-spliced group II intron. *Science 320*(5872), 77–82.

[237] Wang, J. (2010) Inclusion of weak high-resolution X-ray data for improvement of a group II intron structure. *Acta Crystallogr. D Biol. Crystallogr. 66*(9), 988–1000.

[238] Chan, R. T., Robart, A. R., Rajashankar, K. R., Pyle, A. M., and Toor, N. (2012) Crystal structure of a group II intron in the pre-catalytic state. *Nat. Struct. Mol. Biol 19*(5), 555–557.

[239] de Lencastre, A., Hamill, S., and Pyle, A. M. (2005) A single active-site region for a group II intron. *Nat. Struct. Mol. Biol 12*(7), 626–627.

[240] Hamill, S., and Pyle, A. M. (2006) The receptor for branch-site docking within a group II intron active site. *Mol. Cell 23*(6), 831–840.

[241] Dai, L., Chai, D., Gu, S.-Q., Gabel, J., Noskov, S. Y., Blocker, F. J., Lambowitz, A. M., and Zimmerly, S. (2008) A three-dimensional model of a group II intron RNA and its interaction with the intron-encoded reverse transcriptase. *Mol. Cell 30*(4), 472–485.

[242] Qin, P. Z., and Pyle, A. M. (1998) The architectural organization and mechanistic function of group II intron structural elements. *Curr. Opin. Struct. Biol. 8*(3), 301–308.

[243] Fedorova, O., and Zingler, N. (2007) Group II introns: structure, folding and splicing mechanism. *Biol. Chem. 388*(7), 665–678.

[244] Pyle, A. M., Fedorova, O., and Waldsich, C. (2007) Folding of group II introns: a model system for large, multidomain RNAs? *Trends Biochem. Sci. 32*(3), 138–145.

[245] Jacquier, A., and Michel, F. (1987) Multiple exon-binding sites in class II self-splicing introns. *Cell 50*(1), 17–29.

[246] Chanfreau, G., and Jacquier, A. (1996) An RNA conformational change between the two chemical steps of group II self-splicing. *EMBO J. 15*(13), 3466–76.

[247] Costa, M., Deme, E., Jacquier, A., and Michel, F. (1997) Multiple tertiary interactions involving domain II of group II self-splicing introns. *J. Mol. Biol. 267*(3), 520–536.

[248] Podar, M., Dib-Hajj, S. D., and Perlman, P. S. (1995) A UV-induced, Mg(2+)-dependent crosslink traps an active form of domain 3 of a self-splicing group II intron. *RNA 1*(8), 828–840.

[249] Jestin, J.-L., Deme, E., and Jacquier, A. (1997) Identification of structural elements critical for inter-domain interactions in a group II self-splicing intron. *EMBO J. 16*(10), 2945–2954.

[250] Fedorova, O., and Pyle, A. M. (2008) A conserved element that stabilizes the

group II intron active site. *RNA 14*(6), 1048–1056.

[251] Zimmerly, S., Hausner, G., and Wu, X. (2001) Phylogenetic relationships among group II intron ORFs. *Nucleic Acids Res. 29*(5), 1238–1250.

[252] Jarrell, K. A., Dietrich, R. C., and Perlman, P. S. (1988) Group II intron domain 5 facilitates a trans-splicing reaction. *Mol. Cell. Biol. 8*(6), 2361–2366.

[253] Bachl, J., and Schmelzer, C. (1990) Effect of deletions at structural domains of group II intron bI1 on self-splicing in vitro. *J. Mol. Biol. 212*(1), 113–125.

[254] Koch, J. L., Boulanger, S. C., Dib-Hajj, S. D., Hebbar, S. K., and Perlman, P. S. (1992) Group II introns deleted for multiple substructures retain self-splicing activity. *Mol. Cell. Biol. 12*(5), 1950–1958.

[255] Franzen, J. S., Zhang, M., and Peebles, C. L. (1993) Kinetic analysis of the 5' splice junction hydrolysis of a group II intron promoted by domain 5. *Nucleic Acids Res. 21*(3), 627–634.

[256] Sigel, R. K. O., Vaidya, A., and Pyle, A. M. (2000) Metal ion binding sites in a group II intron core. *Nat. Struct. Biol. 7*(12), 1111–1116.

[257] Swisher, J. F., Su, L. J., Brenowitz, M., Anderson, V. E., and Pyle, A. M. (2002) Productive folding to the native state by a group II intron ribozyme. *J. Mol. Biol. 315*(3), 297–310.

[258] Waldsich, C., and Pyle, A. M. (2007) A folding control element for tertiary collapse of a group II intron ribozyme. *Nat. Struct. Mol. Biol 14*(1), 37–44.

[259] Pyle, A. M., and Green, J. B. (1994) Building a kinetic framework for group II intron ribozyme activity: quantitation of interdomain binding and reaction rate. *Biochemistry 33*(9), 2716–2725.

[260] Michels, W. J., and Pyle, A. M. (1995) Conversion of a group II intron into a new multiple-turnover ribozyme that selectively cleaves oligonucleotides: elucidation of reaction mechanism and structure/function relationships. *Biochemistry 34*(9), 2965–2977.

[261] van der Veen, R., Kwakman, J. H., and Grivell, L. A. (1987) Mutations at the lariat acceptor site allow self-splicing of a group II intron without lariat formation. *EMBO J. 6*(12), 3827–3831.

[262] Daniels, D. L., Michels, W. J., and Pyle, A. M. (1996) Two competing pathways for self-splicing by group II introns: a quantitative analysis of in vitro reaction rates and products. *J. Mol. Biol. 256*(1), 31–49.

[263] Peebles, C. L., Zhang, M., Perlman, P. S., and Franzen, J. S. (1995) Catalytically critical nucleotide in domain 5 of a group II intron. *Proc. Natl. Acad. Sci. U. S. A. 92*(10), 4422–4426.

[264] Michel, F., and Ferat, J.-L. (1995) Structure and activities of group II introns. *Annu. Rev. Biochem. 64*, 435–461.

[265] Boulanger, S. C., Belcher, S. M., Schmidt, U., Dib-Hajj, S. D., Schmidt, T., and Perlman, P. S. (1995) Studies of point mutants define three essential paired nucleotides in the domain 5 substructure of a group II intron. *Mol. Cell. Biol. 15*(8), 4479–4488.

[266] Boudvillain, M., and Pyle, A. M. (1998) Defining functional groups, core structural features and inter-domain tertiary contacts essential for group II intron self-splicing: a NAIM analysis. *EMBO J. 17*(23), 7091–7104.

[267] Konforti, B. B., Abramovitz, D. L., Duarte, C. M., Karpeisky, A., Beigelman, L., and Pyle, A. M. (1998) Ribozyme catalysis from the major groove of group II intron domain 5. *Mol. Cell 1*(3), 433–441.

[268] Schmidt, U., Podar, M., Stahl, U., and Perlman, P. S. (1996) Mutations of the two-nucleotide bulge of D5 of a group II intron block splicing in vitro and in vivo: phenotypes and suppressor mutations. *RNA 2*(11), 1161–1172.

[269] Chanfreau, G., and Jacquier, A. (1994) Catalytic site components common to both splicing steps of a group II intron. *Science 266*(5189), 1383–1387.

[270] Abramovitz, D. L., Friedman, R. A., and Pyle, A. M. (1996) Catalytic role of 2'-hydroxyl groups within a group II intron active site. *Science 271*(5254), 1410–1413.

[271] Costa, M., and Michel, F. (1995) Frequent use of the same tertiary motif by self-folding RNAs. *EMBO J. 14*(6), 1276–1285.

[272] Konforti, B. B., Liu, Q., and Marie Pyle, A. (1998) A map of the binding site for catalytic domain 5 in the core of a group II intron ribozyme. *EMBO J. 17*(23), 7105–7117.

[273] Fedorova, O., and Pyle, A. M. (2005) Linking the group II intron catalytic domains: tertiary contacts and structural features of domain 3. *EMBO J. 24*(22), 3906–3916.

[274] Boudvillain, M., de Lencastre, A., and Pyle, A. M. (2000) A tertiary interaction that links active-site domains to the 5' splice site of a group II intron. *Nature 406*(6793), 315–318.

[275] Podar, M., Zhuo, J., Zhang, M., Franzen, J. S., Perlman, P. S., and Peebles, C. L. (1998) Domain 5 binds near a highly conserved dinucleotide in the joiner linking domains 2 and 3 of a group II intron. *RNA 4*(2), 151 –166.

[276] Jacquier, A., and Michel, F. (1990) Base-pairing interactions involving the 5' and 3'-terminal nucleotides of group II self-splicing introns. *J. Mol. Biol. 213*(3), 437–447.

[277] Costa, M., Michel, F., and Westhof, E. (2000) A three-dimensional perspective on exon binding by a group II self-splicing intron. *EMBO J. 19*(18), 5007–5018.

[278] Swisher, J. F., Duarte, C. M., Su, L. J., and Pyle, A. M. (2001) Visualizing the solvent-inaccessible core of a group II intron ribozyme. *EMBO J. 20*(8), 2051–2061.

[279] Henriksen, N., Davis, D., and Cheatham III, T. (2012) Molecular dynamics re-refinement of two different small RNA loop structures using the original NMR data suggest a common structure. *J. Biomol. NMR* 1–19.

[280] Zhang, L., and Doudna, J. A. (2002) Structural insights into group II intron catalysis and branch-site selection. *Science 295*(5562), 2084–2088.

[281] Sigel, R. K. O., Sashital, D. G., Abramovitz, D. L., Palmer III, A. G., Butcher, S. E., and Pyle, A. M. (2004) Solution structure of domain 5 of a group II intron ribozyme reveals a new RNA motif. *Nat. Struct. Mol. Biol 11*(2), 187–192.

[282] Seetharaman, M., Eldho, N. V., Padgett, R. A., and Dayie, K. T. (2006) Structure of a self-splicing group II intron catalytic effector domain 5: Parallels with spliceosomal U6 RNA. *RNA 12*(2), 235–247.

[283] Gumbs, O. H., Padgett, R. A., and Dayie, K. T. (2006) Fluorescence and solution NMR study of the active site of a 160-kDa group II intron ribozyme. *RNA 12*(9), 1693–1707.

[284] Toor, N., Keating, K. S., and Pyle, A. M. (2009) Structural insights into RNA splicing. *Curr. Opin. Struct. Biol. 19*(3), 260–266.

[285] Gordon, P. M., Sontheimer, E. J., and Piccirilli, J. A. (2000) Kinetic characterization of the second step of group II intron splicing: role of metal ions and the cleavage site 2'-OH in catalysis. *Biochemistry 39*(42), 12939–12952.

[286] Chen, Y., Eldho, N. V., Dayie, K. T., and Carey, P. R. (2010) Probing adenine rings and backbone linkages using base specific isotope-edited Raman spectroscopy: application to group II intron ribozyme domain V. *Biochemistry 49*(16), 3427–3435.

[287] Valadkhan, S. (2010) Role of the snRNAs in spliceosomal active site. *RNA Biol. 7*(3), 345–353.

[288] Yu, Y. T., Maroney, P. A., Darzynkiwicz, E., and Nilsen, T. W. (1995) U6 snRNA function in nuclear pre-mRNA splicing: a phosphorothioate interference analysis of the U6 phosphate backbone. *RNA 1*(1), 46–54.

[289] Villa, T., Pleiss, J. A., and Guthrie, C. (2002) Spliceosomal snRNAs: Mg2+-dependent chemistry at the catalytic core? *Cell 109*(2), 149–152.

[290] Butcher, S. E., and Brow, D. A. (2005) Towards understanding the catalytic core structure of the spliceosome. *Biochem. Soc. Trans. 33*(Pt 3), 447–449.

[291] Newman, A., and Norman, C. (1992) U5 snRNA interacts with exon sequences at 5' and 3' splice sites. *Cell 68*(4), 743–754.

[292] Sashital, D. G., Cornilescu, G., McManus, C. J., Brow, D. A., and Butcher, S. E. (2004) U2-U6 RNA folding reveals a group II intron-like domain and a four-helix junction. *Nat. Struct. Mol. Biol 11*(12), 1237–1242.

[293] Madhani, H. D., and Guthrie, C. (1992) A novel base-pairing interaction between U2 and U6 snRNAs suggests a mechanism for the catalytic activation of the spliceosome. *Cell 71*(5), 803–817.

[294] Valadkhan, S., and Manley, J. L. (2001) Splicing-related catalysis by protein-free snRNAs. *Nature 413*(6857), 701–707.

[295] Valadkhan, S., and Manley, J. L. (2003) Characterization of the catalytic activity of U2 and U6 snRNAs. *RNA 9*(7), 892–904.

[296] Valadkhan, S., Mohammadi, A., Jaladat, Y., and Geisler, S. (2009) Protein-free small nuclear RNAs catalyze a two-step splicing reaction. *Proc. Natl. Acad. Sci. 106*(29), 11901–11906.

[297] Huppler, A., Nikstad, L. J., Allmann, A. M., Brow, D. A., and Butcher, S. E. (2002) Metal binding and base ionization in the U6 RNA intramolecular stem-loop structure. *Nat. Struct. Biol. 9*(6), 431–435.

[298] Reiter, N. J., Blad, H., Abildgaard, F., and Butcher, S. E. (2004) Dynamics in the U6 RNA intramolecular stem-loop: a base flipping conformational change. *Biochemistry 43*(43), 13739–13747.

[299] Blad, H., Reiter, N. J., Abildgaard, F., Markley, J. L., and Butcher, S. E. (2005) Dynamics and metal ion binding in the U6 RNA intramolecular stem-loop as analyzed by NMR. *J. Mol. Biol. 353*(3), 540–555.

[300] Hilliker, A. K., and Staley, J. P. (2004) Multiple functions for the invariant AGC triad of U6 snRNA. *RNA 10*(6), 921–928.

[301] Valadkhan, S., and Manley, J. L. (2002) Intrinsic metal binding by a spliceosomal RNA. *Nat. Struct. Mol. Biol 9*(7), 498–499.

[302] Sontheimer, E. J., Sun, S., and Piccirilli, J. A. (1997) Metal ion catalysis during splicing of premessenger RNA. *Nature 388*(6644), 801–805.

[303] Yean, S. L., Wuenschell, G., Termini, J., and Lin, R. J. (2000) Metal-ion coordination by U6 small nuclear RNA contributes to catalysis in the spliceosome. *Nature 408*(6814), 881–884.

[304] Sun, J. S., and Manley, J. L. (1995) A novel U2-U6 snRNA structure is necessary for mammalian mRNA splicing. *Genes Dev. 9*(7), 843–854.

[305] Valadkhan, S., Mohammadi, A., Wachtel, C., and Manley, J. L. (2007) Protein-free spliceosomal snRNAs catalyze a reaction that resembles the first step of splicing. *RNA 13*(12), 2300–2311.

[306] Su, L. J., Brenowitz, M., and Pyle, A. M. (2003) An alternative route for the folding of large RNAs: Apparent two-state folding by a group II intron ribozyme. *J. Mol. Biol. 334*(4), 639–652.

[307] Steiner, M., Karunatilaka, K. S., Sigel, R. K. O., and Rueda, D. (2008) Single-

molecule studies of group II intron ribozymes. *Proc. Natl. Acad. Sci. U. S. A. 105*(37), 13853–13858.

[308] Su, L. J., Waldsich, C., and Pyle, A. M. (2005) An obligate intermediate along the slow folding pathway of a group II intron ribozyme. *Nucleic Acids Res. 33*(21), 6674–6687.

[309] Qin, P. Z., and Pyle, A. M. (1997) Stopped-flow fluorescence spectroscopy of a group II intron ribozyme reveals that domain 1 is an independent folding unit with a requirement for specific Mg^{2+} ions in the tertiary structure. *Biochemistry 36*(16), 4718–4730.

[310] Fedorova, O., Waldsich, C., and Pyle, A. M. (2007) Group II intron folding under near-physiological conditions: collapsing to the near-native state. *J. Mol. Biol. 366*(4), 1099–1114.

[311] Solem, A., Zingler, N., and Pyle, A. M. (2006) A DEAD protein that activates intron self-splicing without unwinding RNA. *Mol. Cell 24*(4), 611–617.

[312] Halls, C., Mohr, S., Del Campo, M., Yang, Q., Jankowsky, E., and Lambowitz, A. M. (2007) Involvement of DEAD-box proteins in group I and group II intron splicing. Biochemical characterization of Mss116p, ATP hydrolysis-dependent and -independent mechanisms, and general RNA chaperone activity. *J. Mol. Biol. 365*(3), 835–855.

[313] Séraphin, B., Simon, M., Boulet, A., and Faye, G. (1989) Mitochondrial splicing requires a protein from a novel helicase family. *Nature 337*(6202), 84–87.

[314] Potratz, J. P., Del Campo, M., Wolf, R. Z., Lambowitz, A. M., and Russell, R. (2011) ATP-dependent roles of the DEAD-box protein Mss116p in group II intron splicing in vitro and in vivo. *J. Mol. Biol. 411*(3), 661–679.

[315] Karunatilaka, K. S., Solem, A., Pyle, A. M., and Rueda, D. (2010) Single-molecule analysis of Mss116-mediated group II intron folding. *Nature 467*(7318), 935–939.

[316] Fedorova, O., Solem, A., and Pyle, A. M. (2010) Protein-facilitated folding of group II intron ribozymes. *J. Mol. Biol. 397*(3), 799–813.

[317] Zingler, N., Solem, A., and Pyle, A. M. (2010) Dual roles for the Mss116 cofactor during splicing of the ai5γ group II intron. *Nucleic Acids Res. 38*(19), 6602 – 6609.

[318] Del Campo, M., Tijerina, P., Bhaskaran, H., Mohr, S., Yang, Q., Jankowsky, E., Russell, R., and Lambowitz, A. M. (2007) Do DEAD-box proteins promote group II intron splicing without unwinding RNA? *Mol. Cell 28*(1), 159–166.

[319] Fedorova, O., and Pyle, A. M. (2012) The brace for a growing scaffold: Mss116 protein promotes RNA folding by stabilizing an early assembly intermediate. *J. Mol. Biol. 422*(3), 347–365.

[320] Waldsich, C., and Pyle, A. M. (2008) A kinetic intermediate that regulates proper folding of a group II intron RNA. *J. Mol. Biol. 375*(2), 572–580.

[321] Liebeg, A., Mayer, O., and Waldsich, C. (2010) DEAD-box protein facilitated RNA folding in vivo. *RNA Biol. 7*(6), 803–811.

[322] Berman, H. M., Westbrook, J., Feng, Z., Gilliland, G., Bhat, T. N., Weissig, H., Shindyalov, I. N., and Bourne, P. E. (2000) The Protein Data Bank. *Nucleic Acids Res. 28*(1), 235–242.

[323] Ebrahimi, M., Rossi, P., Rogers, C., and Harbison, G. S. (2001) Dependence of 13C NMR chemical shifts on conformations of rna nucleosides and nucleotides. *J. Magn. Reson. 150*(1), 1–9.

[324] Ohlenschläger, O., Haumann, S., Ramachandran, R., and Görlach, M. (2008) Conformational signatures of 13C chemical shifts in RNA ribose. *J. Biomol. NMR 42*(2), 139–142.

[325] Ghose, R., Marino, J. P., Wiberg, K. B., and Prestegard, J. H. (1994) Dependence of 13C chemical shifts on glycosidic torsional angles in ribonucleic acids. *J. Am. Chem. Soc. 116*(19), 8827–8828.

[326] Neuhaus, D., and Williamson, M. *The Nuclear Overhauser Effect in structural and conformational analysis*; VCH Publishers, 1989.

[327] Foloppe, N., Hartmann, B., Nilsson, L., and MacKerell Jr., A. D. (2002) Intrinsic conformational energetics associated with the glycosyl torsion in DNA: A quantum mechanical study. *Biophys. J. 82*(3), 1554–1569.

[328] Kao, C., Zheng, M., and Rüdisser, S. (1999) A simple and efficient method to reduce nontemplated nucleotide addition at the 3 terminus of RNAs transcribed by T7 RNA polymerase. *RNA 5*(9), 1268 –1272.

[329] Kao, C., Rüdisser, S., and Zheng, M. (2001) A simple and efficient method to transcribe RNAs with reduced 3' heterogeneity. *Methods 23*(3), 201–205.

[330] Phillips, R. C., Phillips, S. J., George, P., and Rutman, R. J. (1963) Potentiometric studies of the secondary phosphate ionizations of AMP, ADP, and ATP, and calculations of thermodynamic data for the hydrolysis reactions. *Biochemistry 2*(3), 501–508.

[331] Büchner, P., Maurer, W., and Rüterjans, H. (1978) Nitrogen-15 nuclear magnetic resonance spectroscopy of 15N-labeled nucleotides. *J. Magn. Reson. 29*(1), 45–63.

[332] Wang, C., Gao, H., Gaffney, B. L., and Jones, R. A. (1991) Nitrogen-15-labeled oligodeoxynucleotides. 3. Protonation of the adenine N1 in the A.C and A.G mispairs of the duplexes d[CG(15N1)AGAATTCCCG]2 and d[CGGGAATTC(15N1)ACG]2. *J. Am. Chem. Soc. 113*(14), 5486–5488.

[333] Live, D., Radhakrishnan, I., Misra, V., and Patel, D. (1991) Characterization of protonated cytidine in oligonucleotides by 15N NMR studies at natural abundance. *J. Am. Chem. Soc. 113*(12), 4687–4688.

[334] Legault, P., and Pardi, A. (1994) In situ probing of adenine protonation in RNA by 13C NMR. *J. Am. Chem. Soc. 116*(18), 8390–8391.

[335] Benoit, R. L., and Frechette, M. (1986) 1H and 13C nuclear magnetic resonance and ultraviolet studies of the protonation of cytosine, uracil, thymine, and related compounds. *Can. J. Chem. 64*(12), 2348–2352.

[336] Lukavsky, P. J., Kim, I., Otto, G. A., and Puglisi, J. D. (2003) Structure of HCV IRES domain II determined by NMR. *Nat. Struct. Biol. 10*(12), 1033–1038.

[337] Kennard, O., and Hunter, W. N. (1991) Single-crystal x-ray diffraction studies of oligonucleotides and oligonucleotide–drug complexes. *Angew. Chem., Int. Ed. 30*(10), 1254–1277.

[338] Li, Y., Zon, G., and Wilson, W. D. (1991) Thermodynamics of DNA duplexes with adjacent G·A mismatches. *Biochemistry 30*(30), 7566–7572.

[339] Leonard, G. A., McAuley-Hecht, K. E., Ebel, S., Lough, D. M., Brown, T., and Hunter, W. N. (1994) Crystal and molecular structure of r(CGCGAAUUAGCG): an RNA duplex containing two G(anti).A(anti) base pairs. *Structure 2*(6), 483–494.

[340] Gao, X., and Patel, D. J. (1988) G(syn).A(anti) mismatch formation in DNA dodecamers at acidic pH: pH-dependent conformational transition of G·A mispairs detected by proton NMR. *J. Am. Chem. Soc. 110*(15), 5178–5182.

[341] Carbonnaux, C., Van der Marel, G. A., Van Boom, J. H., Guschlbauer, W., and Fazakerley, G. V. (1991) Solution structure of an oncogenic DNA duplex containing a G·A mismatch. *Biochemistry 30*(22), 5449–5458.

[342] Pan, B., Mitra, S. N., and Sundaralingam, M. (1999) Crystal structure of an RNA 16-mer duplex R(GCAGAGUUAAAUCUGC)2 with nonadjacent

G(syn).A+(anti) mispairs. *Biochemistry 38*(9), 2826–2831.

[343] Lane, A. N., Jenkins, T. C., Brown, D. J., and Brown, T. (1991) N.m.r. determination of the solution conformation and dynamics of the A.G mismatch in the d(CGCAAATTGGCG)2 dodecamer. *Biochem. J. 279*(Pt 1), 269–281.

[344] Millar, D. P. (1996) Fluorescence studies of DNA and RNA structure and dynamics. *Curr. Opin. Struct. Biol. 6*(3), 322–326.

[345] Ward, D. C., Reich, E., and Stryer, L. (1969) Fluorescence studies of nucleotides and polynucleotides I. Formycin, 2-aminopurine riboside, 2,6-diaminopurine riboside, and their derivatives. *J. Biol. Chem. 244*(5), 1228–1237.

[346] Steiner, M. Ph.D. thesis, University of Zurich: Zurich, 2008.

[347] Schmitz, M., and Tinoco, I. (2000) Solution structure and metal-ion binding of the P4 element from bacterial RNase P RNA. *RNA 6*(9), 1212–1225.

[348] Schwieters, C. D., Kuszewski, J. J., Tjandra, N., and Clore, G. M. (2003) The Xplor-NIH NMR molecular structure determination package. *J. Magn. Reson. 160*(1), 65–73.

[349] Schwieters, C. D., Kuszewski, J. J., and Marius Clore, G. (2006) Using Xplor–NIH for NMR molecular structure determination. *Prog. Nucl. Magn. Reson. Spectrosc. 48*(1), 47–62.

[350] Davis, A. R., Kirkpatrick, C. C., and Znosko, B. M. (2011) Structural characterization of naturally occurring RNA single mismatches. *Nucleic Acids Res. 39*(3), 1081–1094.

[351] Patel, D. J., Kozlowski, S. A., Ikuta, S., and Itakura, K. (1984) Deoxyguanosine-deoxyadenosine pairing in the d(C-G-A-G-A-A-T-T-C-G-C-G) duplex: conformation and dynamics at and adjacent to the dG.dA mismatch site. *Biochemistry 23*(14), 3207–3217.

[352] Brown, T., Hunter, W. N., Kneale, G., and Kennard, O. (1986) Molecular structure of the G·A base pair in DNA and its implications for the mechanism of transversion mutations. *Proc. Natl. Acad. Sci. 83*(8), 2402–2406.

[353] Webster, G. D., Sanderson, M. R., Skelly, J. V., Neidle, S., Swann, P. F., Li, B. F., and Tickle, I. J. (1990) Crystal Structure and Sequence-Dependent Conformation of the A.G Mispaired Oligonucleotide d(CGCAAGCTGGCG). *Proc. Natl. Acad. Sci. 87*(17), 6693–6697.

[354] Brown, T., Leonard, G. A., Booth, E. D., and Kneale, G. (1990) Influence of pH on the conformation and stability of mismatch base-pairs in DNA. *J. Mol. Biol. 212*(3), 437–440.

[355] Leonard, G. A., Booth, E. D., and Brown, T. (1990) Structural and thermodynamic studies on the adenine.guanine mismatch in B-DNA. *Nucleic Acids Res. 18*(19), 5617–5623.

[356] Morse, S. E., and Draper, D. E. (1995) Purine–purine mismatches in RNA helices: evidence for protonated G·A pairs and next-nearest neighbor effects. *Nucleic Acids Res. 23*(2), 302–306.

[357] Jucker, F. M., Heus, H. A., Yip, P. F., Moors, E. H., and Pardi, A. (1996) A network of heterogeneous hydrogen bonds in GNRA tetraloops. *J. Mol. Biol. 264*(5), 968–980.

[358] Hammond, N. B., Tolbert, B. S., Kierzek, R., Turner, D. H., and Kennedy, S. D. (2010) RNA internal loops with tandem AG pairs: the structure of the 5'GAGU/3'UGAG loop can be dramatically different from others, including 5'AAGU/3'UGAA. *Biochemistry 49*(27), 5817–5827.

[359] Clore, G. M., and Kuszewski, J. (2003) Improving the accuracy of NMR structures of RNA by means of conformational database potentials of mean force as assessed by complete dipolar coupling cross-validation. *J. Am. Chem. Soc.*

125(6), 1518–1525.

[360] Hoffmann, B., Mitchell, G. T., Gendron, P., Major, F., Andersen, A. A., Collins, R. A., and Legault, P. (2003) NMR structure of the active conformation of the Varkud satellite ribozyme cleavage site. *Proc. Natl. Acad. Sci. U. S. A. 100*(12), 7003–7008.

[361] Olivier, C., Poirier, G., Gendron, P., Boisgontier, A., Major, F., and Chartrand, P. (2005) Identification of a conserved RNA motif essential for She2p recognition and mRNA localization to the yeast bud. *Mol. Cell. Biol. 25*(11), 4752–4766.

[362] Lisi, V., and Major, F. (2007) A comparative analysis of the triloops in all high-resolution RNA structures reveals sequence–structure relationships. *RNA 13*(9), 1537–1545.

[363] Zimmermann, G. R., Jenison, R. D., Wick, C. L., Simorre, J. P., and Pardi, A. (1997) Interlocking structural motifs mediate molecular discrimination by a theophylline-binding RNA. *Nat. Struct. Biol. 4*(8), 644–649.

[364] Locker, N., Easton, L. E., and Lukavsky, P. J. (2007) HCV and CSFV IRES domain II mediate eIF2 release during 80S ribosome assembly. *EMBO J. 26*(3), 795–805.

[365] Zuo, X., Wang, J., Yu, P., Eyler, D., Xu, H., Starich, M. R., Tiede, D. M., Simon, A. E., Kasprzak, W., Schwieters, C. D., Shapiro, B. A., and Wang, Y.-X. (2010) Solution structure of the cap-independent translational enhancer and ribosome-binding element in the 3' UTR of turnip crinkle virus. *Proc. Natl. Acad. Sci. 107*(4), 1385–1390.

[366] Khanna, M., Wu, H., Johansson, C., Caizergues-Ferrer, M., and Feigon, J. (2006) Structural study of the H/ACA snoRNP components Nop10p and the 3' hairpin of U65 snoRNA. *RNA 12*(1), 40–52.

[367] Gendron, P., Lemieux, S., and Major, F. (2001) Quantitative analysis of nucleic acid three-dimensional structures. *J. Mol. Biol. 308*(5), 919–936.

[368] Lemieux, S., and Major, F. (2002) RNA canonical and non-canonical base pairing types: a recognition method and complete repertoire. *Nucleic Acids Res. 30*(19), 4250–4263.

[369] Schmeing, T. M., Voorhees, R. M., Kelley, A. C., and Ramakrishnan, V. (2011) How mutations in tRNA distant from the anticodon affect the fidelity of decoding. *Nat. Struct. Mol. Biol 18*(4), 432–436.

[370] Huang, L., Serganov, A., and Patel, D. J. (2010) Structural insights into ligand recognition by a sensing domain of the cooperative glycine riboswitch. *Mol. Cell 40*(5), 774–786.

[371] Laane, C., Krone, W., Konings, W., Haaker, H., and Veeger, C. (1980) Short-term effect of ammonium chloride on nitrogen fixation by Azotobacter vinelandii and by bacteroids of Rhizobium leguminosarum. *Eur. J. Biochem. 103*(1), 39–46.

[372] Orij, R., Postmus, J., Ter Beek, A., Brul, S., and Smits, G. J. (2009) In vivo measurement of cytosolic and mitochondrial pH using a pH-sensitive GFP derivative in Saccharomyces cerevisiae reveals a relation between intracellular pH and growth. *Microbiology 155*(Pt 1), 268–278.

[373] Latham, M. P., Zimmermann, G. R., and Pardi, A. (2009) NMR chemical exchange as a probe for ligand-binding kinetics in a theophylline-binding RNA aptamer. *J. Am. Chem. Soc. 131*(14), 5052–5053.

[374] Reiter, N. J., Nikstad, L. J., Allmann, A. M., Johnson, R. J., and Butcher, S. E. (2003) Structure of the U6 RNA intramolecular stem–loop harboring an SP-phosphorothioate modification. *RNA 9*(5), 533–542.

[375] Konarska, M. M., Vilardell, J., and Query, C. C. (2006) Repositioning of the

Reaction Intermediate within the Catalytic Center of the Spliceosome. *Mol. Cell 21*(4), 543–553.

[376] Liu, L., Query, C. C., and Konarska, M. M. (2007) Opposing classes of prp8 alleles modulate the transition between the catalytic steps of pre-mRNA splicing. *Nat. Struct. Mol. Biol 14*(6), 519–526.

[377] Mefford, M. A., and Staley, J. P. (2009) Evidence that U2/U6 helix I promotes both catalytic steps of pre-mRNA splicing and rearranges in between these steps. *RNA 15*(7), 1386–1397.

[378] McCallum, S. A., and Pardi, A. (2003) Refined solution structure of the iron-responsive element RNA using residual dipolar couplings. *J. Mol. Biol. 326*(4), 1037–1050.

[379] Schudoma, C., May, P., and Walther, D. (2010) Modeling RNA loops using sequence homology and geometric constraints. *Bioinformatics 26*(13), 1671–1672.

[380] Duarte, C. M., and Pyle, A. M. (1998) Stepping through an RNA structure: A novel approach to conformational analysis. *J. Mol. Biol. 284*(5), 1465–1478.

[381] Addess, K. J., Basilion, J. P., Klausner, R. D., Rouault, T. A., and Pardi, A. (1997) Structure and dynamics of the iron responsive element RNA: implications for binding of the RNA by iron regulatory binding proteins. *J. Mol. Biol. 274*(1), 72–83.

[382] Hall, K. B., and Tang, C. (1998) 13C relaxation and dynamics of the purine bases in the iron responsive element RNA hairpin. *Biochemistry 37*(26), 9323–9332.

[383] Hall, K. B., and Williams, D. J. (2004) Dynamics of the IRE RNA hairpin loop probed by 2-aminopurine fluorescence and stochastic dynamics simulations. *RNA 10*(1), 34–47.

[384] Walden, W. E., Selezneva, A. I., Dupuy, J., Volbeda, A., Fontecilla-Camps, J. C., Theil, E. C., and Volz, K. (2006) Structure of dual function iron regulatory protein 1 complexed with ferritin IRE-RNA. *Science 314*(5807), 1903–1908.

[385] Butcher, S. E., Dieckmann, T., and Feigon, J. (1997) Solution structure of a GAAA tetraloop receptor RNA. *EMBO J. 16*(24), 7490–7499.

[386] Davis, J. H., Tonelli, M., Scott, L. G., Jaeger, L., Williamson, J. R., and Butcher, S. E. (2005) RNA helical packing in solution: NMR structure of a 30 kDa GAAA tetraloop-receptor complex. *J. Mol. Biol. 351*(2), 371–382.

[387] Nozinovic, S., Fürtig, B., Jonker, H. R. A., Richter, C., and Schwalbe, H. (2010) High-resolution NMR structure of an RNA model system: the 14-mer cUUCGg tetraloop hairpin RNA. *Nucleic Acids Res. 38*(2), 683 –694.

[388] Qin, P. Z., Butcher, S. E., Feigon, J., and Hubbell, W. L. (2001) Quantitative analysis of the isolated GAAA tetraloop/receptor interaction in solution: a site-directed spin labeling study. *Biochemistry 40*(23), 6929–6936.

[389] Dingley, A. J., and Grzesiek, S. (1998) Direct observation of hydrogen bonds in nucleic acid base pairs by internucleotide 2JNN couplings. *J. Am. Chem. Soc. 120*(33), 8293–8297.

[390] Heus, H., and Pardi, A. (1991) Structural features that give rise to the unusual stability of RNA hairpins containing GNRA loops. *Science 253*(5016), 191–194.

[391] Erat, M. C., Kovacs, H., and Sigel, R. K. O. (2010) Metal ion-N7 coordination in a ribozyme branch domain by NMR. *J. Inorg. Biochem. 104*(5), 611–613.

[392] Tanaka, Y., and Ono, A. (2008) Nitrogen-15 NMR spectroscopy of N-metallated nucleic acids: insights into 15N NMR parameters and N-metal bonds. *Dalton Trans.* 4965–4974.

[393] Helm, L., and Merbach, A. (1999) Water exchange on metal ions: experiments and simulations. *Coord. Chem. Rev. 187*(1), 151–181.

[394] Sigel, H., Massoud, S. S., and Tribolet, R. (1988) Comparison of the metal

ion coordinating properties of tubercidin 5'-monophosphate (7-deaza-AMP) with those of adenosine 5'-monophosphate (AMP) and 1,N6-ethenoadenosine 5'-monophosphate (.epsilon.-AMP). Definite evidence for metal ion-base-backbinding to N-7 and extent of macrochelate formation in M(AMP) and M(.epsilon.-AMP). *J. Am. Chem. Soc. 110*(20), 6857–6865.

[395] Maderia, M., Horton, T. E., and DeRose, V. J. (2000) Metal interactions with a GAAA RNA tetraloop characterized by 31P NMR and phosphorothioate substitutions. *Biochemistry 39*(28), 8193–8200.

[396] Lescoute, A., and Westhof, E. (2006) Topology of three-way junctions in folded RNAs. *RNA 12*(1), 83–93.

[397] Sarver, M., Zirbel, C. L., Stombaugh, J., Mokdad, A., and Leontis, N. B. (2008) FR3D: finding local and composite recurrent structural motifs in RNA 3D structures. *J. Math. Biol. 56*(1-2), 215–252.

[398] Popenda, M., Szachniuk, M., Blazewicz, M., Wasik, S., Burke, E. K., Blazewicz, J., and Adamiak, R. W. (2010) RNA FRABASE 2.0: an advanced web-accessible database with the capacity to search the three-dimensional fragments within RNA structures. *BMC Bioinf. 11*(1), 231.

[399] Geary, C., Chworos, A., and Jaeger, L. (2011) Promoting RNA helical stacking via A-minor junctions. *Nucleic Acids Res. 39*(3), 1066–1080.

[400] Toor, N., Keating, K. S., Fedorova, O., Rajashankar, K., Wang, J., and Pyle, A. M. (2010) Tertiary architecture of the Oceanobacillus iheyensis group II intron. *RNA 16*(1), 57–69.

[401] Tyagi, R., and Mathews, D. H. (2007) Predicting helical coaxial stacking in RNA multibranch loops. *RNA 13*(7), 939–951.

[402] Gallo, S., Furler, M., and Sigel, R. K. O. (2005) In vitro transcription and purification of RNAs of different size. *Chimia 59*(11), 812–816.

[403] Brauer, G., and Handa, B. K. (1964) Heterotype Mischkristalle bei Hexammincobaltjodiden. *Zeitschrift für anorganische und allgemeine Chemie 329*(1), 12–17.

[404] Sigel, H., Zuberbühler, A. D., and Yamauchi, O. (1991) Comments on potentiometric pH titrations and the relationship between pH-meter reading and hydrogen ion concentration. *Anal. Chim. Acta 255*(1), 63–72.

[405] Kankia, B. I., Buckin, V., and Bloomfield, V. A. (2001) Hexamminecobalt(III)-induced condensation of calf thymus DNA: circular dichroism and hydration measurements. *Nucleic Acids Res. 29*(13), 2795–2801.

[406] Sigel, H., and Griesser, R. (2005) Nucleoside 5'-triphosphates: self-association, acid-base, and metal ion-binding properties in solution. *Chem. Soc. Rev. 34*(10), 875–900.

[407] Glasoe, P. K., and Long, F. A. (1960) Use of glass electrodes to measure acidities in deuterium oxide. *J. Phys. Chem. 64*(1), 188–190.

[408] Markley, J. L., Bax, A., Arata, Y., Hilbers, C. W., Kaptein, R., Sykes, B. D., Wright, P. E., and Wüthrich, K. (1998) Recommendations for the presentation of NMR structures of proteins and nucleic acids. *J. Mol. Biol. 280*(5), 933–952.

[409] Schanda, P., and Brutscher, B. (2005) Very fast two-dimensional NMR spectroscopy for real-time investigation of dynamic events in proteins on the time scale of seconds. *J. Am. Chem. Soc. 127*(22), 8014–8015.

[410] Davis, A. L., Keeler, J., Laue, E. D., and Moskau, D. (1992) Experiments for recording pure-absorption heteronuclear correlation spectra using pulsed field gradients. *J. Magn. Reson. 98*(1), 207–216.

[411] Kellogg, G. W. (1992) Proton-detected hetero-TOCSY experiments with application to nucleic acids. *Journal of Magnetic Resonance (1969) 98*(1), 176–182.

[412] Zwahlen, C., Legault, P., Vincent, S. J. F., Greenblatt, J., Konrat, R., and Kay, L. E. (1997) Methods for Measurement of Intermolecular NOEs by Multinuclear NMR Spectroscopy: Application to a Bacteriophage λ N-Peptide/boxB RNA Complex. *J. Am. Chem. Soc. 119*(29), 6711–6721.

[413] Tjandra, N., and Bax, A. (1997) Measurement of dipolar contributions to 1JCH splittings from magnetic-field dependence of J modulation in two-dimensional NMR spectra. *J. Magn. Reson. 124*(2), 512–515.

[414] Güntert, P., Mumenthaler, C., and Wüthrich, K. (1997) Torsion angle dynamics for NMR structure calculation with the new program DYANA. *J. Mol. Biol. 273*(1), 283–298.

[415] Fürtig, B., Richter, C., Bermel, W., and Schwalbe, H. (2004) New NMR experiments for RNA nucleobase resonance assignment and chemical shift analysis of an RNA UUCG tetraloop. *J. Biomol. NMR 28*(1), 69–79.

[416] Aboul-ela, F., Karn, J., and Varani, G. (1995) The structure of the Human Immunodeficiency Virus type-1 TAR RNA reveals principles of RNA recognition by Tat protein. *J. Mol. Biol. 253*(2), 313–332.

[417] Varani, G., Cheong, C., and Tinoco, I. (1991) Structure of an unusually stable RNA hairpin. *Biochemistry 30*(13), 3280–3289.

[418] Gorenstein, D. *Phosphorus-31 NMR : principles and applications*; Academic Press: Orlando Fla., 1984.

[419] Brünger, A. T., Adams, P. D., Clore, G. M., DeLano, W. L., Gros, P., Grosse-Kunstleve, R. W., Jiang, J. S., Kuszewski, J., Nilges, M., Pannu, N. S., Read, R. J., Rice, L. M., Simonson, T., and Warren, G. L. (1998) Crystallography & NMR system: A new software suite for macromolecular structure determination. *Acta Crystallogr. D Biol. Crystallogr. 54*(Pt 5), 905–921.

[420] Brünger, A. T. (2007) Version 1.2 of the Crystallography and NMR system. *Nat. Protoc. 2*(11), 2728–2733.

[421] Clore, G. M., Gronenborn, A. M., and Tjandra, N. (1998) Direct structure refinement against residual dipolar couplings in the presence of rhombicity of unknown magnitude. *J. Magn. Reson. 131*(1), 159–162.

[422] Koradi, R., Billeter, M., and Wüthrich, K. (1996) MOLMOL: a program for display and analysis of macromolecular structures. *J. Mol. Graph. 14*(1), 51–55, 29–32.

[423] Baker, N. A., Sept, D., Joseph, S., Holst, M. J., and McCammon, J. A. (2001) Electrostatics of nanosystems: application to microtubules and the ribosome. *Proc. Natl. Acad. Sci. U. S. A. 98*(18), 10037–10041.

[424] Draper, D. E., Gluick, T. C., and Michael L. Johnson, G. K. A. Melting studies of RNA unfolding and RNA-ligand interactions. In *Energetics of Biological Macromolecules*; Academic Press, 1995; Vol. Volume 259, pp 281–305.

[425] Marky, L. A., and Breslauer, K. J. (1987) Calculating thermodynamic data for transitions of any molecularity from equilibrium melting curves. *Biopolymers 26*(9), 1601–1620.

[426] Handloser, C. S., Chakrabarty, M. R., and Mosher, M. W. (1973) Experimental determination of pKa values by use of NMR chemical shift. *J. Chem. Educ. 50*(7), 510.

[427] Wilcox, J. L., Ahluwalia, A. K., and Bevilacqua, P. C. (2011) Charged nucleobases and their potential for RNA catalysis. *Acc. Chem. Res. 44*(12), 1270–1279.

[428] Wijmenga, S. S., Mooren, M. M., and Hilbers, C. W. NMR of nucleic acids; from spectrum to structure. In *NMR of macromolecules: a practical approach*; The Practical Approach Series; Oxford University Press, 1993; pp 217–280.

Curriculum Vitae

Personal data

Name	Maria Pechlaner
Date and Place of Birth	July 28, 1984 in Innsbruck, Austria
Nationality	Austrian

Education

1990 - 1994	Primary school in Innsbruck
1994 - 2002	Grammar school Adolf-Pichler-Platz, Innsbruck
June 2002	Graduation (“Matura”)
Oct 2002 - Apr 2008	Studies of Molecular Biology at the University of Vienna, Austria
March 2007 - March 2008	Diploma Thesis at the Department of Biomolecular Structural Chemistry, University of Vienna, under supervision of Prof. Kristina Djinovic-Carugo Title: “*Radiation damage in metalloproteins – a scavenger approach*”
Nov 2008 - Dec 2012	Ph.D. thesis with Prof. Dr. Roland K.O. Sigel at the Institute of Inorganic Chemistry, University of Zurich

Courses

July 29 - Aug 5, 2011	EMBO Practical Course “*Structure, dynamics and function of biomacromolecules by solution NMR*”, BNMRZ Munich, Germany
Jan 8 - 13, 2012	10th NCCR Practical Course “*Introduction to Biomolecular Modelling*”, Kandersteg, Switzerland

Teaching experience

Oct - Dec 2009	Teaching assistant in the general chemistry lab course at the Departement of Chemistry and Biochemistry of the University of Zurich
Nov 2010	Teaching assistant in the project laboratory course at the Institute of Inorganic chemistry of the University of Zurich

Publications

- *Can soaked-in scavengers protect metalloprotein active sites from reduction during data collection?* S. Macedo, M. Pechlaner, W. Schmid, M. Weik, K. Sato, Ch. Dennison, K. Djinovic-Carugo (2009) J. Synchrotron Rad. 16, 191-204.
- *Characterization of metal ion-nucleic acid interactions in solution* M. Pechlaner and R. K. O. Sigel. (2012) Met. Ions Life Sci. 10, 1-42.
- *The structural stabilization of the κ three-way junction by Mg(II) represents the first step in the folding of a group II intron* D. Donghi, M. Pechlaner, C. Finazzo, B. Knobloch and R. K. O. Sigel. (2012) Nucleic Acids Res. *In press, doi:10.1093/nar/gks1179.*
- *Acid-base equilibria near neutrality in the catalytic triad and the bulge in domain 5 of a bacterial group II intron.* M. Pechlaner, D. Donghi, V. Zelenay, R. K. O. Sigel *To be submitted shortly.*
- *Structure and dynamics of the domain 5 hairpin of a bacterial group II intron based on molecular dynamics simulations and NMR data* M. Pechlaner, J. Dolenc, R. K. O. Sigel, W. F. van Gunsteren *In preparation.*

Poster presentations

- Fall meeting of the Swiss Chemical Society 2009 (Lausanne, Switzerland): *Structural investigation and metal ion titration of a bacterial group II intron catalytic core*, M. Pechlaner, D. Donghi, V. Zelenay and R. K. O. Sigel
- Swiss RNA Workshop 2010 (Bern, Switzerland): *Structural investigation of the catalytic core of a bacterial group II intron*, M. Pechlaner, D. Donghi, V. Zelenay and R.K.O. Sigel
- 10th European Biological Inorganic Chemistry Conference (EuroBIC) 2010 (Thessaloniki, Greece): *Structure and metal ion binding of the central domain of a bacterial group II intron*, M. Pechlaner, D. Donghi and R.K.O. Sigel
- Fall meeting of the Swiss Chemical Society 2010 (Zürich, Switzerland): *Solution structure of the catalytic core of a bacterial group II intron*, M. Pechlaner, D. Donghi and R.K.O. Sigel

- Fall meeting of the Swiss Chemical Society 2011 (Lausanne, Switzerland): *Mg^{2+} binding sites in the κ-ζ region of group II intron Sc.ai5γ*, M. Pechlaner, D. Donghi and R.K.O. Sigel
- Swiss RNA Workshop 2012 (Bern, Switzerland): *Stabilization of a group II intron core structure by di- and trivalent metal ions*, M. Pechlaner, D. Donghi and R. K. O. Sigel
- XXVth International Conference on Magnetic Resonance in Biological Systems (ICMRBS) 2012 (Lyon, France): *An RNA Three-Way Junction at the Group II Intron Core is Stabilized by Mg(II)*, M. Pechlaner, D. Donghi and R. K. O. Sigel
- Fall meeting of the Swiss Chemical Society 2012 (Zürich, Switzerland): *Solution NMR studies demonstrate the stabilization of the group II intron core by di- and trivalent metal ions*, M. Pechlaner, D. Donghi and R. K. O. Sigel